Memory & Cognition
1998, 26 (6), 1270-1281

Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences

HILDUR E. H. SCHILLING, KEITH RAYNER, and JAMES I. CHUMBLEY
University of Massachusetts, Amherst, Massachusetts

Performance on three different tasks was compared: naming, lexical decision, and reading (with eye fixation times on a target word measured). We examined the word frequency effect for a common set of words for each task and each subject. Naming and reading (particularly gaze duration) yielded similar frequency effects for the target words. The frequency effect found in lexical decision was greater than that found in naming and in eye fixation times. In all tasks, there was a correlation between the frequency effect and average response time. In general, the results suggest that both the naming and the lexical decision tasks yield data about word recognition processes that are consistent with effects found in eye fixations during silent reading.

A major goal of research using naming and lexical decision tasks has been to understand word recognition. While this is an important objective, many researchers hope that the results obtained with such tasks generalize beyond the scope of word recognition studies to reading per se. Given that so much research has relied on these tasks, it is not surprising that some investigations have attempted to understand the basic components of each task (Balota & Chumbley, 1985; Chumbley & Balota, 1984). It is also not surprising that several studies have compared these two tasks with each other and with categorization tasks (Lewellen, Goldinger, Pisoni, & Greene, 1993; Monsell, Doyle, & Haggard, 1989). Lexical decision performance has also recently been compared with neuroimaging profiles obtained from functional magnetic resonance imaging (Pugh et al., 1997). However, no studies have directly compared performance on naming and lexical decision tasks with eye fixation times during reading. Some studies have indirectly compared eye fixation times with either naming or lexical decision by using the same stimulus materials and have yielded similar context effects (Altarriba, Kroll, Sholl, & Rayner, 1996; Schustack, Ehrlich, & Rayner, 1987) or phonological coding effects (Pollatsek, Lesch, Morris, & Rayner, 1992) with naming and eye fixation times. In contrast, Everatt and Underwood (1994) found little correlation between eye fixation times and lexical decision times.

In the present study, we compared naming, lexical decision, and eye fixation times for high-frequency (HF) and low-frequency (LF) target words. Given that the word frequency effect (i.e., LF words are responded to more slowly than HF words) is quite robust, we reasoned that examining the nature of the effect across the three tasks would be more informative than simply correlating overall response time (RT) with fixation time (as was done by Everatt & Underwood, 1994). The research reported here had three goals. First, we were interested in a direct comparison of the three tasks. In particular, it is important to determine how well the widely used naming and lexical decision tasks correlate with eye fixation times since the latter type of data are obtained while subjects are actually reading and have been shown to reflect cognitive processes in reading (Rayner & Sereno, 1994; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Second, we were interested in the extent to which the word frequency effect was consistent across the three tasks. Prior research has demonstrated a word frequency effect in naming (Balota & Chumbley, 1985; Forster & Chambers, 1973), lexical decision (Chumbley & Balota, 1984; Whaley, 1978), and eye fixation times (Inhoff & Rayner, 1986; Just & Carpenter, 1980; Rayner, 1977; Rayner & Duffy, 1986; Rayner & Fischer, 1996; Rayner & Raney, 1996; Rayner, Sereno, & Raney, 1996). However, there has not been a direct comparison across the three tasks. Third, we were interested in examining individual differences across the three tasks.

Although no prior research has compared all three tasks with each other, some research has examined the relationship between lexical decision and naming within the context of the frequency effect.[1] Specifically, the frequency effect has been found to be smaller in naming than

This research was supported by NSF Grant DBS-9121375 and NIH Grant HD26765. K.R. was also supported by a research scientist award from the NIMH (MH01255). Portions of the data were presented at the joint meeting of the Experimental Psychology Society and the Canadian Society for Brain, Behaviour, and Cognitive Science, Toronto, July 1993. We thank Arnold Well for his assistance with data analyses, and Charles Clifton, Morton A. Gernsbacher, Wayne Murray, and three anonymous reviewers for their helpful comments on an earlier version of the article. Correspondence should be addressed to K. Rayner, Department of Psychology, University of Massachusetts, Amherst, MA 01003 (e-mail: rayner@psych.umass.edu).

in lexical decision (Balota & Chumbley, 1984; Besner & McCann, 1987; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Hudson & Bergman, 1985; Paap, McDonald, Schvaneveldt, & Noel, 1987; Richardson, 1976; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). One explanation for this finding is that words displayed visually can be pronounced using spelling-to-sound correspondence rules; pronouncing words on the basis of phonology may minimize lexical access processes. That is, if word frequency primarily affects processing at a lexical level, then naming should be less susceptible to the effect of frequency than is lexical decision, which requires access to lexical representations to determine whether a letter string is a word. Furthermore, response latency in the lexical decision task does not necessarily reflect only the time that it takes to access a word in the lexicon: The size of the frequency effect may be inflated by decision processes (Balota & Chumbley, 1984). In deciding whether a letter string is a word, subjects may consider its familiarity (Gernsbacher, 1984). Compared with HF words, LF words are more similar in familiarity to nonwords. Thus, it may take longer to respond to LF than to HF words since letter strings are less familiar and are more difficult to discriminate from nonwords. Because longer response latencies for LF words may be due to difficulty in discrimination rather than a difference in the time it takes to access words in the lexicon, lexical decision may not be as informative as naming for studying the effect of word frequency on lexical access.

Although lexical decision involves a decision component not required in reading, naming requires oral production of words, which is also not necessary for silent reading. Balota and Chumbley (1985) found that when subjects were given 1,400 msec to access a word before producing it, there was still a frequency effect. This suggests that word frequency influences a production stage of naming that may not be important in reading.

Given that many researchers utilizing naming and lexical decision studies intend to generalize their results to reading, it is important to examine the relationship between RTs in these two tasks and fixation times in reading. However, there is some controversy over exactly what kinds of processes are reflected in the two standard fixation time measures: first-fixation duration and gaze duration. First-fixation duration is the duration of the first fixation on a target word. If only one fixation on the word is made, that fixation is used to compute the mean fixation time. But, if two or more fixations are made on the word, only the first one is used. Gaze duration is the total fixation time on a target word prior to moving to another word; it does not include the duration of regressive fixations made to the target word after fixation on another word. Inhoff (1984) initially argued that first-fixation duration reflects primarily lexical processes whereas gaze duration reflects lexical processes and text integration processes. However, since many studies have yielded similar results across the two measures, Rayner and Pollatsek (1987) argued that the two measures reflect the same processes. In the present study, both measures were compared with naming and lexical decision RTs.

Although one of our primary interests in this study was comparing the frequency effect across tasks, it is the case that the frequency effect has been used to study individual differences in reading skill; skilled readers exhibit a smaller frequency effect than less skilled readers (Seidenberg, 1985a; Seidenberg et al., 1984; Waters & Seidenberg, 1985). Waters, Seidenberg, and Bruck (1984) found that less skilled readers read LF strange words (e.g., *aisle*, *tsar*) more slowly than LF regular words, whereas skilled readers showed no differences in the two types of words. They attributed this difference in the regularity effect to the hypothesis that HF words are recognized on a visual basis whereas LF words that take longer to recognize are "sounded out" and hence are affected by regularity. Thus, when recognizing a LF word, skilled readers may process the whole word on a visual basis, while less skilled readers may depend on the phonological route to "sound it out." Therefore, a smaller frequency effect is exhibited by skilled readers than less skilled readers whose LF RTs are influenced by regularity. Seidenberg (1985b) claimed that the regularity effect skilled readers show for LF words is more reliable in naming than in lexical decision. In the present article, we examined individual differences in relation to naming, lexical decision, and eye fixation times.

METHOD

Subjects

Forty-eight students at the University of Massachusetts at Amherst participated in the experiment. They were either given course credit or paid $5 for participation. All of the subjects were native speakers of English with normal vision.

Design

Three pairs of tasks were used with 16 subjects randomly assigned to each task pair: naming and lexical decision; naming and reading; and lexical decision and reading. The task order was counterbalanced so that for each task pair, 8 subjects completed a particular task first and 8 subjects completed it second. At least 1 week intervened between the two tasks.[2]

Materials

For each task, word or sentence order was randomized independently for each subject. Twenty-four HF and 24 LF words were selected from Balota and Chumbley (1985). According to the Francis and Kučera (1987) norms, the HF words had counts greater than 46 per million (mean rating = 141 occurrences per million), whereas the LF words had counts less than 4 per million (mean rating = 1.92 per million). The HF and LF words were matched on word length (which ranged from 6 to 9 letters). They were presented individually in the naming and lexical decision tasks, and each word was embedded in a sentence in the reading task (see Appendix A for the sentences). The sentence context prior to the target word was intended to be relatively neutral and to not strongly constrain the target word (see Appendix B for predictability ratings).[3] For each task, appropriate practice items were used. For lexical decision, 48 nonwords (created by changing one letter in a word) were added.

Apparatus for Naming and Lexical Decision Tasks

The subjects were tested in a sound-deadened room. An intercom enabled the experimenter to communicate with the subjects. Word stimuli were generated by an IMSAI VIO Video Interface Board and were presented as uppercase letters on a Zenith ZVM 1230A video monitor approximately 50 cm in front of the subjects and simultaneously on a second monitor for the experimenter. Three letters spanned approximately 1° of visual angle. The video monitor was controlled by a Northstar Horizon computer. A voicekey connected to the computer was used for the naming task: Microswitches connected to the computer were used for the lexical decision task. The computer recorded response latency in milliseconds.

Procedure for the Naming Task

Subjects were told that a word would be presented on each trial and that they should pronounce it as quickly as possible without stuttering or mispronouncing it. There were six blocks of trials. The first two blocks were practice blocks of 12 trials to acquaint the subjects with the procedure. The subsequent four test blocks of 14 trials each contained 2 practice trials followed by 12 test trials. On each trial, a 500-Hz warning tone sounded for 250 msec. The word was displayed on the monitor 250 msec following tone offset. When the voicekey detected a response, the word was erased from the screen. The experimenter scored the trial as a correct pronunciation of the word or as an "error" if the word was mispronounced or if an irrelevant noise (such as a cough) triggered the voicekey. After the response was recorded, there was a 2-sec intertrial interval before the tone signaling the next trial. After each block of trials, the average RT and percentage of trials in which a correct response was recorded for that block was displayed on the subjects' monitor. Ten seconds later, a message on the screen signaled that the subjects could press a button to continue the experiment.

Procedure for the Lexical Decision Task

Subjects were told that on each trial a letter string would be presented, and that they should decide as quickly as possible if it was a word. There were 10 blocks of trials. The first two blocks were practice blocks of 12 trials to acquaint subjects with the procedure. The subsequent eight test blocks of 14 trials each contained 2 practice trials followed by 12 test trials (six words and six nonwords). On each trial, a 500-Hz warning tone sounded for 250 msec. The letter string was displayed on the monitor 250 msec following tone offset. The apparatus for lexical decision was the same as that used in the naming task except that subjects responded by pulling microswitch-mounted levers placed on the table in front of them. The lever by the dominant hand was marked "yes" and the one by the nondominant hand was marked "no." Subjects responded by pulling the lever marked "yes" if the string was a word and by pulling the lever marked "no" if the string was not a word. When the response lever was pulled, the letter string was erased from the screen. If the subject made a mistake, the message ERROR was displayed on the screen, and the subject pressed a button to continue the study. After the response was made, there was a 2-sec intertrial interval before the tone signaling the next trial. Feedback was presented at the end of each block for the lexical decision task.

Procedure and Apparatus for the Reading Task

The sentences were displayed on a SONY Trinitron 1302 monitor. All but the first letter of each sentence was in lowercase. Eye movements were recorded with a Fourward Technologies Dual-Purkinje Eyetracker, which has a resolution of 10′ of arc. Viewing was binocular, with eye movements monitored from the right eye, which was 62 cm from the monitor. Four characters corresponded to 1° of visual angle. The eyetracker and the monitor were connected to an Epson Equity III computer, which controlled the experiment and recorded the duration and location of each fixation. When a subject arrived for the experiment, a bite bar (which was used to eliminate head movements) was prepared and the eyetracking system was calibrated. Subjects were instructed to silently read the sentences and to pull a lever after they read each sentence.

There were 48 test sentences preceded by five practice sentences to acquaint subjects with the procedure. Before presentation of each sentence, five boxes were presented on the video screen. The subjects looked at the center box and then at the first box on the left. When the subject looked at the left box, the experimenter pressed a button to erase the boxes from the screen and to display a sentence. The subjects read the sentence, and as soon as they pulled the lever, the sentence was erased from the screen. On 12 trials, after the sentence was read, the word QUESTION appeared, followed by a statement pertaining to the sentence for that trial. These questions were to ensure that the subjects were comprehending the sentences. The subjects pressed the right lever if the statement was true and pressed the left lever if the statement was false. If subjects made a mistake, the message ERROR was displayed on the screen for 1.5 sec. The subjects answered the questions correctly over 95% of the time. Fixations less than 100 msec were eliminated from the data analyses unless there was a fixation on an adjacent letter. Such short fixations are typically assumed to be due to oculomotor programming factors (R. E. Morrison, 1984) and to not reflect lexical processing. Less than 2% of the data were excluded because of track losses of the eyetracking system.

RESULTS

We will refer to the dependent variables (naming latency, lexical decision latency, first-fixation duration, gaze duration, and single-fixation duration[4]) as performance measures. An alpha level of .05 was used for all analyses. Four different aspects of the data were examined: (1) basic analyses of the means of the performance measures, the word frequency effect for each measure, and the correct response rates or fixation and refixation probabilities; (2) comparisons of frequency effect sizes across measures; (3) analyses of the stability of individual subjects' frequency effects across measures; and (4) analyses of the relationship between a subject's frequency effect and mean RT.

Basic Analyses

RTs and frequency effects for each response measure are displayed in Table 1. The data for the naming and lexical decision measures are based on correct responses to the target words. For naming, subjects made correct responses to 98% of the HF words and to 97% of the LF words. For lexical decision, more correct responses were made to HF targets (97%) than to LF targets (89%). RT was 698 msec for the 96% of the nonword targets responded to correctly. For reading, a fixation was scored if it was on the target word or the space preceding it. Subjects fixated on HF words 89% of the time and on LF words 93% of the time; they refixated HF words 11% of the time and LF words 23% of the time.

Because each subject performed only two of the three tasks, a 2 (task pair) × 2 (frequency) mixed analysis of variance (ANOVA) on the data from each of the measures examined the effects of task pairing and order in which tasks were performed. There were no main effects of task pair (all Fs < 2.53, all ps $> .12$), indicating that average performance in each task did not depend on the

Table 1
Response Times (RTs), Frequency Effects (FEs), and Their Standard Errors (*SEMs*) for the Performance Measures (in Milliseconds)*

	Performance Measure				
	Naming Latency	Lexical Decision Latency	First-Fixation Latency	Gaze Duration	Single-Fixation Duration
Mean RT	546	594	248	291	260
SEM	11.3	15.9	6.3	9.6	7.3
Low-frequency RT	578	671	265	324	286
SEM	20.7	30.8	11.1	17.4	14.2
High-frequency RT	514	522	230	257	237
SEM	12.6	17.4	8.1	10.9	9.5
Frequency effect	64	149	35	67	49
SEM	7.8	13.9	5.7	7.2	7.9
Slope of FE function†	−1.73	−3.85	−0.89	−1.78	−1.26
SE of slope	0.21	0.36	0.14	0.18	0.19

*Since the analyses in the table are based on the data of all 32 subjects who performed a task, each subject contributed data to performance measures from two tasks. †The FE function regresses RT (in milliseconds) on 40+10*log(word frequency + 1).

other task with which it was paired. There were no main effects of order (all Fs < 1), which indicates that average RT was not affected by order of task.

There was a main effect of word frequency for the different performance measures using both subjects (F_1) and items (F_2) as random variables within each measure. For naming, mean RT for LF words (578 msec) was significantly greater than that for HF words (514 msec) [$F_1(1,28) = 71.13, MS_e = 951; F_2(1,46) = 44.98, MS_e =$ 1,140]. For lexical decision, the mean RT for LF words (671 msec) was significantly greater than that for HF words (522 msec) [$F_1(1,28) = 124.22$, $MS_e = 2{,}855$; $F_2(1,46) = 98.6$, $MS_e = 2{,}940$]. For reading, the mean first-fixation duration on LF words (265 msec) was significantly greater than that on HF words (230 msec) [$F_1(1,28) = 37.02, MS_e = 574; F_2(1,46) = 50.18, MS_e =$ 314], and the gaze duration on LF words (324 msec) was significantly greater than that on HF words (257 msec) [$F_1(1,28) = 95.53, MS_e = 758; F_2(1,46) = 55.89, MS_e =$ 1,003].[5] The only significant interaction (all other Fs $<$ 2.48, ps $> .12$) was in the lexical decision task between order and frequency [$F_1(1,28) = 4.36$, $MS_e = 2{,}855$; $F_2(1,46) = 5.28$, $MS_e = 2{,}912$], which was due to the fact that the average frequency effect of 177 msec when lexical decision was the first task was reduced to 121 msec when preceded by either the naming or the reading task. Since there were no main effects of task pair or order, the following analyses address response and fixation times that are averaged across these variables.

Frequency Effects Across Tasks

Since the design of the experiment involved each subject participating in two tasks, tests for any differences in the frequency effect between performance measures could not be performed on the data as shown in Table 1. Instead, two complementary analyses using within-subjects and between-subjects comparisons were conducted to determine whether the magnitude of the frequency effect varied across the different performance measures.

Within-subjects comparisons were conducted to determine whether a subject's frequency effect on one performance measure differed from his/her frequency effect on a second measure. For subjects who performed the naming and lexical decision tasks, the average frequency effect in lexical decision (142 msec) was larger than that in naming (75 msec) [$t(15) = 4.3$]. For subjects who performed the naming and reading tasks, the average frequency effect in first fixation (41 msec) was similar to that in naming (55 msec); the two measures did not differ significantly [$t(15) = 1.25, p > .05$]. The average frequency effect in gaze duration (78 msec) was larger than that in naming (55 msec) [$t(15) = 2.32$]. For subjects who performed lexical decision and reading, the average frequency effect in lexical decision (156 msec) was larger than that in both first fixation (32 msec) [$t(15) = 5.79$] and gaze duration (56 msec) [$t(15) = 5.49$]. These within-subjects analyses indicate that the size of the word frequency effect in lexical decision was larger than that in both naming and reading. Also, the size of the word frequency effect for naming was similar to the size of the effect for first-fixation duration and smaller than that for gaze duration.

The second comparison of the sizes of frequency effects used between-subjects comparisons. Since there were three tasks, half of the subjects performing a given task performed one of the remaining two tasks, and the other subjects performed the third task. For example, half of the subjects performing the naming task also engaged in the lexical decision task and the other half engaged in the reading task. Using the two groups of subjects with a common task (e.g., naming), the sizes of the frequency effects in the other two tasks (reading and lexical decision) could be compared without concern about differential effects of the paired tasks. It should be noted, however, that each subject contributed to two of the between-subjects analyses that will now be reported.

For the 32 subjects who participated in the naming task, the 16 who also engaged in lexical decision produced a frequency effect of 142 msec for lexical decision. The 16 who had reading as their other task produced significantly smaller frequency effects for both first fixation (40 msec) [$t(30) = 4.98$] and gaze duration (78 msec) [$t(30) = 3.02$]. Of the 32 subjects who had reading as one of their tasks, the 16 who also performed lexical decision produced a 156-msec frequency effect for lexical decision. This effect was significantly [$t(30) =$ 4.45] larger than the 55-msec frequency effect produced by the 16 subjects who had naming as their other task. Finally, of the subjects who participated in lexical decision, the 16 who also engaged in reading produced a 32-msec frequency effect for first fixation and a 56-msec

Table 2
Correlations of Subjects' Response Times (RTs) for the Two Tasks

	Task Pair				
	Naming and Lexical Decision	Naming and First Fixation	Lexical Decision and First Fixation	Naming and Gaze Duration	Lexical Decision and Gaze Duration
Mean RT (by subjects)	.610	.666	.571	.774	.602
Mean RT (by words)	.833	.504	.654	.528	.711
High-frequency RT	.486	.309	.376	.532	.480
Low-frequency RT	.657	.674	.616	.808	.644
Frequency effect	.575	.238	.152	.542	.542

Note—Analyses with first-fixation and gaze duration as performance measures for reading. With 16 subjects in a group, an $r(14)$ that equals $\pm.4259$ is significantly different from zero. For the mean RT, correlations shown are based on subjects' mean RTs (top row) and on the means of the individual words (second row).

effect for gaze duration. Only the first-fixation effect was significantly smaller [$t(3) = 2.89$] than the 75-msec effect produced by the 16 subjects who had naming as their second task. The effects for gaze duration and naming were not significantly different [$t(30) = 1.21, p > .05$].

The within- and between-subjects comparisons of reading with naming yielded slightly different results. As noted, the within-subjects frequency effects for naming, first fixation, and gaze duration were 55, 41, and 78 msec, respectively. In contrast, the between-subjects analyses yielded effects of 75, 32, and 56 msec. Although the size of the frequency effect for first fixation was about the same absolute size in both analyses and was the smallest frequency effect in both, the absolute and relative sizes of the frequency effect in naming and gaze duration varied between analyses. We do not have a ready explanation for this except to note that, as seen in Table 1, the standard errors of the frequency effects for naming and gaze duration are approximately a third larger than the standard error of the frequency effect for first fixation. In any case, a conservative conclusion that is consistent with the effects displayed in Table 1 is that naming and gaze duration yield frequency effects of similar size, which are larger than the frequency effect for first fixation.

The last statistic in Table 1, the slope of the frequency effect function, is an alternative measure to the frequency effect. In the present experiment, we followed the typical procedure and selected sets of HF and LF words; that is, we constructed a dichotomous distribution of word frequencies. The frequency effect is the difference in average RT to these word sets. On occasion it is inconvenient or impossible to select words in this way, but it may still be important to assess the effects of word frequency. In these cases, the procedure adopted by Balota and Chumbley (1984) provides a solution: the frequency effect function, which regresses RT on $40 + 10*\log(\text{word frequency} + 1)$. Not only does this procedure permit assessment of frequency effects when the words have not been selected for this purpose, it also permits other interesting comparisons that will be described below. In any case, as seen in Table 1, the slope of the frequency effect function mirrors the frequency effects calculated in the standard way.

Stability of Individual Differences in Reading Speed and Frequency Effect

One clue to the presence of common processes and structures involved in the two tasks is the extent to which subjects are at the same relative performance level in the tasks as indexed by the correlation of a response measure for pairs of tasks. Table 2 presents data indicating that there indeed are significant correlations in performance[6] for all three pairs of tasks on four different measures: overall RT, RT for HF words, RT for LF words, and the frequency effect.

The two measures of most interest (which are not necessarily correlated themselves; see below) are overall RT and the frequency effect. A correlation in overall RTs for the two tasks might be expected for a number of reasons, but there are also reasons to expect the correlations to be small since the tasks logically differ in several components. For example, lexical decision requires a decision about which response to make, naming requires that a pronunciation be determined, and reading requires that the meaning of a word be integrated with that of other words in the sentence. A similar consideration applies to the frequency effect since it has been demonstrated that manipulations of difficulty in arriving at a decision, in pronouncing a word, and in integrating a meaning can affect it. To the extent that the frequency effect is produced by different task properties, we would expect small correlations. On the other hand, if there is a common component to the frequency effects in all tasks and there are reliable individual differences in the size of this component, there should be a high correlation between subjects' frequency effects across tasks.

The relative size of an individual's frequency effect tended to be the same in the two tasks performed when the performance measures were naming, lexical decision, and gaze duration (all of the correlations of the frequency effects were significantly greater than zero). This finding suggests that a process sensitive to word frequency (or some covariate of it) and common to the tasks affected performance measures in all three tasks. Interestingly, the frequency effect in first fixation was not significantly correlated with the frequency effect in lexical decision or naming.

Relationship Between the Frequency Effect and RT

Seidenberg (1985b) suggested that an individual's reading ability, as measured by RT measures, is related to the size of the frequency effect exhibited by the individual. He tested this hypothesis by using a median split based on overall naming time to divide subjects into skilled and less skilled reading groups and then tested for a difference between the two groups in the size of the frequency effect. Since overall RT in our three tasks, although correlated, would not necessarily yield the same division of subjects into ability groups and a median split ignores differences within ability levels, a more sensitive measure of the relationship between overall RT and the frequency effect was adopted, the correlation between mean overall RT and the frequency effect. As demonstrated in Appendix C, the numerator of this correlation is the difference between the RT variance for HF and LF words. The denominator is the square root of the product of the variance of the difference and the variance of the sum. Thus, to the extent that there is greater variability across subjects in RTs to LF words than to HF words, the correlation will be positive. Equal variances will produce a correlation of zero and, in the unlikely event that RTs to HF words exhibit greater variability, the correlation will be negative.

The correlations were calculated on the data displayed in Table 1. Naming latency, lexical decision latency, first-fixation duration, gaze duration, and single-fixation duration were all analyzed. The results of correlating a subject's frequency effect with his/her mean RT are displayed in the first row of Table 3. Since 32 subjects contributed to each measure in Table 3, correlations that are greater than or equal to .349 are significant with $p < .05$ by a two-tailed test and those greater than or equal to .449 are significant with $p < .01$. Thus, the frequency effect a subject produces is significantly correlated with the subject's mean RT for all five performance measures. This is equivalent to the result observed by Seidenberg (1985b) and extends his result to performance measures other than naming latency.

As seen in Table 1, the slope of the frequency effect function depends, as does the frequency effect, upon the performance measure under consideration, and the variation across measures is quite large. In addition, it is worth noting that the size of the frequency effect is not correlated with the mean RT for the group. Measures with nearly the same mean RT can exhibit frequency effects that differ by a factor of two. It is of some interest, then, to determine whether the relationship between a subject's frequency effect and mean RT also varies across measures. In order to make this determination, a more informative statistic than the correlation is required. The slopes of the relationship between word frequency and RT provide such a statistic.

Since there is a positive correlation between the frequency effect and subject mean RT, the slope of the regression of the frequency effect on subjects' mean RTs is the millisecond increase in the frequency effect with each millisecond increase in mean RT. The second and third lines of Table 3 display the slopes and their standard errors for each performance measure. All of the slopes are far greater than 2 *SE*s and thus are clearly different from zero. What is more striking, however, is that the slopes are roughly the same. A precise statistical test is not available because of the way subjects performed pairs of tasks. It can be noted, however, that the smallest possible standard error of the difference between the means of two sets of scores (which occurs when there is a perfect positive correlation between the scores) is a standard error of the mean based on the pooled variances. Using this fact, noting that the smallest standard error is 0.084, and enforcing a rough cutoff of 2 *SE*s, all but the slope for first fixation fall within 2 *SE*s of the other slopes. Thus, within the resolving power of our study and acknowledging the danger in reasoning based on null effects, the frequency effect changes by about 0.5 msec with each 1-msec change in mean RT for all performance measures, and this is in stark contrast to the highly different (using the same criteria) frequency effects observed for the performance measures in Table 1.

Seidenberg (1985b) noted that when the frequency effect is different at different levels of some variable (including reading speed), most of the difference is concentrated in changed RTs to LF words. Rows 4–9 of Table 3 present analyses of the relationship between RT and subject mean RT for both LF and HF words. Although the results displayed in these rows follow from the fact that subjects varied more in their RTs to LF words than they did to HF words (see Appendix C), they

Table 3
Correlations and Slopes of Functions Relating Subject Mean Response Times (SMRTs) and Measures of the Effect of Word Frequency

	Performance Measure				
	Naming Latency	Lexical Decision Latency	First-Fixation Duration	Gaze Duration	Single-Fixation Duration
Relationship between subject's frequency effect and SMRT					
Correlation	.747	.703	.419	.662	.434
Slope of function	0.515	0.614	0.382	0.499	0.466
SE of slope	0.084	0.113	0.151	0.103	0.177
Relationship between subject's low-frequency RT and SMRT					
Correlation	.984	.972	.947	.976	.921
Slope of function	1.252	1.310	1.172	1.235	1.243
SE of slope	0.041	0.058	0.072	0.050	0.096
Relationship between subject's high-frequency RT and SMRT					
Correlation	.954	.915	.878	.929	.856
Slope of function	0.738	0.697	0.790	0.736	0.777
SE of slope	0.042	0.056	0.079	0.053	0.086
Relationship between slope of subject's frequency effect function and SMRT*					
Correlation	.748	.714	.368	.667	.375
Slope of function	0.014	0.016	0.008	0.012	0.010
SE of slope	0.002	0.003	0.004	0.003	0.004

*The frequency effect function regresses RT in milliseconds on 40 + 10*log(word frequency+1). The correlations and slopes have been multiplied by −1 to avoid double negatives.

are of interest. Namely, there are three important findings in these rows. First, the slopes for LF words are greater than the slopes for HF words (since, of course, the slope for the frequency effect is simply the difference between the LF and HF slopes; see Appendix C). RT increases by about 1.25 msec for each 1-msec change in mean RT for LF words but increases by only about 0.75 msec for each 1-msec change in mean RT for HF words. This indicates that our data are in accord with those of Seidenberg (1985b). Second, all of the slopes are clearly different from zero since they are much more than 2 *SE*s greater than zero. This implies that subjects differ in their RTs to HF words as well as in their RTs to LF words. Finally, note that there is a striking similarity in the slopes across performance measures for both LF and HF words. This similarity reinforces the suggestion that whatever is producing the frequency effect in one measure has a similar effect on all measures.

The bottom three rows of Table 3 present one final approach to investigating the relationship between a subject's mean RT and the frequency effect produced by the subject. The analysis uses the concept of the frequency effect function introduced in describing the bottom row of Table 1. A frequency effect function was calculated for each subject and each measure. The slope of this function provides an estimate of the subject's sensitivity to word frequency. Next, the slopes of the frequency effect functions were regressed against subjects' mean RTs. The slope at the bottom of Table 3 represents the change (in milliseconds) in slope of the frequency effect function with each 1-msec change in subject mean RT. A comparison of the top three rows of Table 3 with the bottom three rows indicates that the two approaches to assessing how sensitivity to word frequency is related to subject mean RT yield quite comparable results. The patterns of data in the top three rows are reproduced in the bottom three. In the present study, there is no advantage of using the frequency effect function approach rather than the frequency effect because the words were selected so that they fell into two groups that had little overlap in word frequency. There are, however, other circumstances where dichotomous distributions cannot be used to examine or control for the effect of word frequency across experimental conditions.

DISCUSSION

The present study had three goals: (1) to compare naming, lexical decision, and reading (eye fixation times) tasks when the comparison for a pair of tasks was made within subjects and for the same words; (2) to examine the stability across tasks of relative performance on overall speed and the size of the word frequency effect; and (3) to examine individual differences.

With respect to how the tasks compare, we found that naming and reading are more similar than the other pairings. Across many analyses, the lexical decision task yielded results somewhat different from those of the other tasks. Specifically, the size of the frequency effect in naming and eye fixation times was similar, while the size of the effect in lexical decision was much larger. In comparing correlations between tasks, naming and reading were more similar than the other pairings. Naming and reading yielded higher correlations than did the other task pairs on RTs to HF and to LF words, and on average RT; the ranking of individual scores for naming was similar to that of scores for reading (eye fixation times).

Lexical decision was more sensitive to repetition effects than were the other tasks since order influenced the frequency effect with this task: The frequency effect in lexical decision as a second task was reduced by prior experience with naming or reading. This finding suggests that there was a repetition effect that persisted from Task 1 to Task 2. As Balota and Chumbley (1984) pointed out, subjects' responses in lexical decision may be influenced by visual familiarity. Since LF words are more similar in familiarity to nonwords than are HF words, they are more difficult to discriminate from nonwords. This discrimination difficulty would yield longer RTs for LF than for HF words. Exposure to LF words in a naming or reading task may increase familiarity for those words. This familiarity boost may persist through subsequent testing. However, familiarity apparently is not as helpful for naming and reading as it is for lexical decision.

Although our data clearly indicate that the naming and eye fixation data were quite similar, the lexical decision data did bear some similarity to the other two tasks. Also, the overall correlation between lexical decision time and eye fixation times was significant. This result contradicts those reported by Everatt and Underwood (1994). Whereas we found significant correlations between lexical decision and first-fixation duration ($r = .571$) and between lexical decision and gaze duration ($r = .602$), Everatt and Underwood reported nonsignificant correlations ($r = .073$ and $.052$). The fact that they did not use the same stimulus materials across their lexical decision and reading tasks may account for the discrepant findings.

As previously found (Balota & Chumbley, 1984; Besner & McCann, 1987; Forster & Chambers, 1973; Frederiksen & Kroll, 1976; Hudson & Bergman, 1985; Paap et al., 1987; Richardson, 1976), the size of the frequency effect was significantly larger in lexical decision than in naming. In contrast, the size of the frequency effect was similar in naming and eye fixations. However, the frequency effect was present in all tasks. Because there was a correlation across tasks among individual subjects, there is something common to be explained. The difference in frequency effect size across tasks can be explained by phonology, oral production, decision processes, movement of one's eye before the word is understood, and text integration processes. The true frequency effect is most likely greater than that found in naming in which LF times are decreased by the use of phonological rules. The true frequency effect is also probably smaller than that found in lexical decision in which

HF words can be responded to on the basis of familiarity. In spite of the many factors that may influence the frequency effect within each task, we found stable individual differences across tasks, which suggests that our findings are robust.

When we examined the performance of individual subjects across tasks, we found that subjects tended to rank themselves consistently in both overall speed and in the size of the frequency effect they produced for all three pairs of tasks. The relatively high correlations ($r \geq .48$) we observed between naming time, lexical decision time, and gaze duration strongly suggest that these tasks incorporate a common lexical access process that is affected similarly by whatever produces differences in the word frequency effect across subjects. The size of the observed effect differed across tasks, presumably because factors such as decision processes (in lexical decision), nonlexical pronunciation processes (in naming), and text integration processes (in gaze duration) modulate the "true" frequency effect. Although some components differed across tasks that might have affected the size of the observed frequency effect, the frequency effect was correlated within subjects across the tasks. The frequency effect in naming was correlated with both that in lexical decision and that in reading as measured by gaze duration. The frequency effect in gaze duration was correlated with that in lexical decision. These correlations can be parsimoniously attributed to a true lexical frequency effect common to all three tasks.

An interesting finding is that the frequency effect in first fixation was not correlated with that in naming or lexical decision. As noted, first-fixation times may be differentially affected by word frequency of the target word. Subjects tend to move their eyes to another letter in a LF word more often than in a HF word, perhaps before having understood it. In this case, there may be additional variability in first-fixation frequency effect differences that does not correspond to that in naming and lexical decision.

In summary, the present study compared performance on naming, lexical decision, and reading tasks. Although the naming data were most like the eye fixation data, it is important to note that the lexical decision data were also significantly correlated with the eye fixation data. Thus, we conclude that both the naming and lexical decision tasks yield data concerning word recognition processes that are consistent with effects found during silent reading.

REFERENCES

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, **24**, 477-492.

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 & Performance*, **10**, 340-357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory & Language*, **24**, 89-106.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.

Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition*, **12**, 590-606.

Everatt, J., & Underwood, G. (1994). Individual differences in reading subprocesses: Relationships between reading ability, lexical access, and eye movement control. *Language & Speech*, **37**, 283-297.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning & Verbal Behavior*, **12**, 627-635.

Francis, W., & Kučera, H. (1987). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 361-379.

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.

Hudson, P. T. W., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory & Language*, **24**, 46-58.

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning & Verbal Behavior*, **23**, 612-624.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, **40**, 431-439.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, **87**, 329-354.

Lewellen, M. J., Goldinger, S. D., Pisoni, D. B., & Greene, B. (1993). Lexical familiarity and processing efficiency: Individual differences in naming, lexical decision, and semantic categorization. *Journal of Experimental Psychology: General*, **122**, 316-330.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 570-590.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General*, **118**, 43-71.

Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 116-133.

Morrison, R. E (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 667-682.

Paap, K. R., McDonald, J. E., Schvaneveldt, R. W., & Noel, R. W. (1987). Frequency and pronounceability in visually presented naming and lexical decision tasks. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 221-243). Hillsdale, NJ: Erlbaum.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in reading. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 148-162.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Fletcher, J. M., Skudlarski, P., Fulbright, R. K., Constable, R. T., Bronen, R. A., Lacadie, C., & Gore, J. C. (1997). Predicting reading performance from neuroimaging profiles: The cerebral basis of phonological effects in printed word identification. *Journal of Experimental Psychology: Human Perception & Performance*, **23**, 299-318.

RAYNER, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, **5**, 443-448.
RAYNER, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J. M. Findlay, R. W. Kentridge, & R. Walker (Eds.), *Eye movement research: Mechanisms, processes, and applications* (pp. 3-22). Amsterdam: North-Holland.
RAYNER, K., & DUFFY, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, **14**, 191-201.
RAYNER, K., & FISCHER, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics*, **58**, 734-747.
RAYNER, K., & POLLATSEK, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 327-362). London: Erlbaum.
RAYNER, K., & RANEY, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, **3**, 245-248.
RAYNER, K., & SERENO, S. C. (1994). Eye movements in reading: Psycholinguistic studies. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 57-81). New York: Academic Press.
RAYNER, K., SERENO, S. C., MORRIS, R. K., SCHMAUDER, A. R., & CLIFTON, C. (1989). Eye movements and on-line comprehension processes. *Language & Cognitive Processes*, **4**, 21-49.
RAYNER, K., SERENO, S. C., & RANEY, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1188-1200.
RICHARDSON, J. T. E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, **67**, 315-325.
SCHUSTACK, M. W., EHRLICH, S. F., & RAYNER, K. (1987). The complexity of contextual facilitation in reading. Local and global inferences. *Journal of Memory & Language*, **26**, 322-340.
SEIDENBERG, M. S. (1985a). Constraining models of word recognition. *Cognition*, **20**, 169-190.
SEIDENBERG, M. S. (1985b). The time-course of phonological code activation in two writing systems. *Cognition*, **19**, 1-30.
SEIDENBERG, M. S., WATERS, G. S., BARNES, M. A., & TANENHAUS, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, **23**, 383-404.
WATERS, G. S., & SEIDENBERG, M. S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition*, **13**, 557-572.
WATERS, G. S., SEIDENBERG, M. S., & BRUCK, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition*, **12**, 293-305.
WHALEY, C. P. (1978). Word-nonword classification time. *Journal of Verbal Learning & Verbal Behavior*, **17**, 143-154.

NOTES

1. There is currently some controversy about whether effects that have been attributed to word frequency are actually due to the correlated age of acquisition factor (see, e.g., C. M. Morrison & Ellis, 1995). Although we agree that the issue is important, it is not directly relevant to the issues addressed in this article. Throughout the article, we refer to "frequency effects," but this can be interpreted as "frequency and/or age of acquisition effects" without affecting any major conclusions. There is also some controversy (Lupker, Brown, & Colombo, 1997) surrounding specifically what kinds of variables affect the frequency effect in naming and lexical decision. Again, although the issue is important, it is not directly relevant to our concerns because our primary goal was to use words that were known to produce a frequency effect in naming and lexical decision and to compare performance on those tasks with reading.

2. Subjects participated in only two tasks, rather than all three, to minimize effects due to repetition of the target words.

3. A subsequent norming study, in which 20 subjects were given the sentence context up to the target word and asked to provide the next word, revealed that some items were more predictable than originally intended. Analyses conducted on the full set of stimuli and on only those that contained low-constraint words yielded identical results (see note 5).

4. Single-fixation duration represents the duration of the fixation on a target word when only one fixation is made on the word. This measure is included in Tables 1 and 3 because some researchers believe that it is more readily interpretable than first-fixation duration and gaze duration measures (see Rayner, 1995, for discussion of these issues) and believe that cases in which readers make only a single fixation on a target word provide the purest measure of processing time. As shown in Table 1, the single-fixation and first-fixation means were quite comparable. An examination of the data revealed that subjects made only a single fixation on the target word (prior to moving to another word) 78% of the time with LF words compared with 67% of the time with HF words.

5. As pointed out in note 3, some of the target words were more predictable than others (see Appendix B). An analysis was conducted on items that were not predictable (words with predictability ratings of .10 or less). For gaze duration, the size of the frequency effect for unpredictable words (69 msec) was similar to that for all items (66 msec). Furthermore, for first-fixation duration, the frequency effect size (36 msec) was equivalent for unpredictable items and all items.

6. Not surprisingly, correlations among the eye fixation measures for the variables shown in Table 2 revealed that first-fixation duration and single-fixation duration were highly correlated ($rs > .90$); gaze duration also correlated with both first-fixation duration and single-fixation duration, with the correlations (rs) ranging from .616 to .916.

APPENDIX A

High-Frequency Words (in Boldface) and the Reading Sentences:

Margie moved into her new **apartment** at the end of the summer.
The principal introduced the new **president** of the junior class.
None of the **students** wanted to have an exam after Spring Break.
Mark told Janet that he would meet her after **baseball** practice.
Bill complained that the **magazine** included more ads than articles.
The angry man called the **senator** to complain about the new tax law.
The policeman demanded to see Jim's **license** and registration.
A strict vegetarian, Jennifer does not eat **chicken** or beef.
Nancy's **kitchen** was infested with carpenter ants and roaches.
The hurricane destroyed houses in the **village** and left many homeless.
Amy told the **teacher** that her dog ate her homework assignment.
Ed was forbidden to attend **college** parties while he was in high school.
The angry man called the **senator** to complain about the new tax law.
Sheri and her **friends** went to Hawaii for their summer vacation.
Mark put too much soap in the washing **machine**, and it overflowed.
The circus tents were crowded with **animals**, clowns, and children.
The bear chased after the **forest** ranger who was carrying honey.
The best place that serves **coffee** and muffins is Dunkin Donuts.
Mr. Jones asked his son to water the **plants** and mow the lawn.
The bride's **mother** cried during the entire wedding ceremony.

APPENDIX A (Continued)

The drunk driver lost control, crashed into a **street** sign and died.
Mary was the only teenager who attended the **square** dance in town.
The burglar broke the **window** and quietly sneaked into the house.
Jimmy was sent to the principal's **office** because he punched Sally.
Most job applications require at least one **letter** of recommendation.

Low-Frequency Words (in Boldface) and the Reading Sentences:

The daredevil was relieved when his **parachute** finally opened.
It is not unusual to see an **armadillo** cross a road in Texas.
Erik took his sick **parakeet** to the veterinarian on Tuesday.
Al stretched before running to avoid pulling a **ligament** or muscle.
The boxer flared his **nostrils** as he entered the boxing ring.
Mary was thrilled to receive a **trinket** from her boyfriend.
Covered with **maggots**, the rug was removed from the smelly dorm room.
Propelled from a submarine, the **torpedo** struck the battleship.
Alfred served baked **haddock** and asparagus to his girlfriend.
The dancer resembled a **gazelle** as he leaped across the stage.
The little girl had dimples in her chin and a **freckle** on her nose.
The beach was covered with **pebbles**, sea shells, and star fish.
The child had a nightmare about being chased by **hornets** and wasps.
Biff dove into the water and retrieved a **scallop** from the ocean floor.
The little girl picked all of the **cashews** out of the trail mix.
The athlete broke his **pelvis** and could not participate in the race.
The game show contestant won a **quartz** watch and a television set.
The careless mailman delivered the **parcel** to the wrong house.
After receiving money, the **beggar** bought cigarettes and a case of beer.
At the science party, people were dressed as **robots** and computers.
The stunning actress wore a black **sequin** dress to the award ceremony.
When the man ran in the blizzard, an **icicle** formed on his beard.
The clumsy assistant dropped a **beaker**, and it shattered on the floor.
When Al's **retina** became inflamed and sore, he visited the eye doctor.

APPENDIX B
Mean Response Times (in Milliseconds) for Naming (N), Lexical Decision (LD), First Fixation (FF), Gaze Duration (GD), Single Fixation (SF), and Predictability Ratings (PR)*

Stimuli	N	LD	FF	GD	SF	PR
APARTMENT	569	555	213	245	217	.75
PRESIDENT	541	526	226	274	236	0
STUDENTS	526	489	217	249	221	.15
BASEBALL	511	540	231	289	241	0
MAGAZINE	505	544	230	269	236	0
SENATOR	540	595	240	281	253	0
LICENSE	494	519	215	291	207	.65
CHICKEN	519	516	250	309	286	0
KITCHEN	500	494	235	251	221	0
VILLAGE	523	509	201	215	200	0
TEACHER	501	492	223	223	219	.15
COLLEGE	506	515	224	248	226	0
FRIENDS	508	527	227	231	224	.35
MACHINE	532	508	226	230	217	.89
ANIMALS	506	506	248	287	250	.10
FOREST	498	510	257	285	245	0
COFFEE	481	481	259	271	253	.15
PLANTS	508	535	241	264	246	.30
MOTHER	481	518	208	216	213	.10
STREET	553	494	225	225	220	0
SQUARE	575	546	251	272	263	0
WINDOW	516	608	231	252	244	.45
OFFICE	451	487	226	237	231	1.00
LETTER	477	511	241	249	241	.05
PARACHUTE	596	652	243	293	257	.15
ARMADILLO	633	829	292	359	307	0
PARAKEET	593	697	273	426	290	0
LIGAMENT	637	763	265	278	264	0
NOSTRILS	611	738	248	275	256	.85
TRINKET	556	737	235	337	263	0
MAGGOTS	552	686	261	331	276	0

APPENDIX B (Continued)

Stimuli	N	LD	FF	GD	SF	PR
TORPEDO	589	741	270	389	280	.30
HADDOCK	566	639	248	322	301	0
GAZELLE	658	742	268	315	282	0
FRECKLE	555	588	243	283	256	.10
PEBBLES	518	575	268	344	298	0
HORNETS	535	698	274	324	277	0
SCALLOP	570	602	275	287	280	0
CASHEWS	578	608	275	316	296	0
PELVIS	569	657	247	307	262	0
QUARTZ	585	613	248	291	236	0
PARCEL	547	606	259	320	279	0
BEGGAR	543	645	260	299	264	0
ROBOTS	517	615	292	337	300	.05
SEQUIN	643	807	317	372	360	0
ICICLE	586	691	272	319	284	.05
BEAKER	535	625	302	332	319	.15
RETINA	597	698	280	347	296	0

*Predictability ratings represent the proportion of subjects (out of 20) who guessed the target word after having read the words that preceded the target word. For each sentence, the first word appeared on the computer screen, and the subject guessed the second word by typing in a response. Next, the correct word was presented, and the subject predicted the subsequent word. This procedure was repeated until the sentence was completed. The order of sentences was randomized independently for each subject.

APPENDIX C

The following formulations are the basis for some observations about the possible relationships between the response times (RTs) for high- and low-frequency words and the frequency effect, as well as the meanings of those relationships.

First, note that

the mean of X is $\mu_X = E[X]$;
the variance of X is $\sigma_X^2 = E[X^2] - \mu_X^2$;
the covariance of X and Y is $\mathrm{Cov}[X,Y] = E[XY] - \mu_X\mu_Y$;
the correlation of X and Y is $\rho_{XY} = \mathrm{Cov}[XY]/\sigma_X\sigma_Y$;
the variance of $X - Y$ is $\sigma_{X-Y}^2 = \sigma_X^2 - 2\rho_{XY}\sigma_X\sigma_Y + \sigma_Y^2 = \sigma_X^2 - 2\mathrm{Cov}[X,Y] + \sigma_Y^2$;
the variance of $X + Y$ is $\sigma_{X+Y}^2 = \sigma_X^2 + 2\mathrm{Cov}[X,Y] + \sigma_Y^2$;
the slope of the regression line of Y on X is $b_{YX} = \mathrm{Cov}[X,Y]/\sigma_X^2$.

From the equation for the correlation and the fact that the largest possible absolute value of the correlation is 1.0, it can be seen that $\mathrm{Cov}[X,Y]$ can never be larger than the larger of the two variances involved.

The following notation is adopted:

L = a subject's RT for low-frequency words;
H = a subject's RT for high-frequency words;
FE = a subject's frequency effect = the difference L − H; and
SMRT = a subject's mean RT = (L + H)/2.

APPENDIX C (Continued)

Using the above formulations and notation:

$$\begin{aligned}
\mathrm{Cov[H,SMRT]} &= \mathrm{Cov[H,(L+H)/2]} \\
&= E[\mathrm{H(L+H)/2}] - \mu_{\mathrm{H}}\mu_{(\mathrm{L+H})/2} \\
&= E[\mathrm{LH}]/2 + E[\mathrm{H}^2]/2 - \mu_{\mathrm{H}}(\mu_{\mathrm{L}}+\mu_{\mathrm{H}})/2 \\
&= (\sigma^2_{\mathrm{H}} + \mathrm{Cov[L,H]})/2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov[L,SMRT]} &= \mathrm{Cov[L,(L+H)/2]} \\
&= E[\mathrm{L(L+H)/2}] - \mu_{\mathrm{L}}\mu_{(\mathrm{L+H})/2} \\
&= E[\mathrm{LH}]/2 + E[\mathrm{L}^2]/2 - \mu_{\mathrm{L}}(\mu_{\mathrm{L}}+\mu_{\mathrm{H}})/2 \\
&= (\sigma^2_{\mathrm{L}} + \mathrm{Cov[L,H]})/2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov[FE,SMRT]} &= \mathrm{Cov[(L-H),(L+H)/2]} \\
&= E[\mathrm{(L-H)(L+H)/2}] - \mu_{\mathrm{L-H}}\mu_{(\mathrm{L+H})/2} \\
&= E[(\mathrm{L}^2-\mathrm{H}^2)/2] - (\mu_{\mathrm{L}}-\mu_{\mathrm{H}})(\mu_{\mathrm{L}}+\mu_{\mathrm{H}})/2 \\
&= E[\mathrm{L}^2]/2 - E[\mathrm{H}^2]/2 - \mu^2_{\mathrm{L}}/2 + \mu^2_{\mathrm{H}}/2 \\
&= (\sigma^2_{\mathrm{L}} - \sigma^2_{\mathrm{H}})/2 \\
&= \mathrm{Cov[L,SMRT]} - \mathrm{Cov[H,SMRT]}
\end{aligned}$$

Therefore:

$$\begin{aligned}
&\rho_{\mathrm{L.SMRT}} = (\sigma^2_{\mathrm{L}} + \mathrm{Cov[L,H]})/2\sigma_{\mathrm{L}}\sigma_{\mathrm{SMRT}} = (\sigma^2_{\mathrm{L}} + \mathrm{Cov[L,H]})/\sigma_{\mathrm{L}}\sigma_{\mathrm{L+H}}; \\
&b_{\mathrm{L.SMRT}} = (\sigma^2_{\mathrm{L}} + \mathrm{Cov[L,H]})/2\sigma^2_{\mathrm{SMRT}} = 2(\sigma^2_{\mathrm{L}} + \mathrm{Cov[L,H]})/\sigma^2_{\mathrm{L+H}}; \\
&\rho_{\mathrm{H.SMRT}} = (\sigma^2_{\mathrm{H}} + \mathrm{Cov[L,H]})/2\sigma_{\mathrm{H}}\sigma_{\mathrm{SMRT}} = (\sigma^2_{\mathrm{H}} + \mathrm{Cov[L,H]})/\sigma_{\mathrm{H}}\sigma_{\mathrm{L+H}}; \\
&b_{\mathrm{H.SMRT}} = (\sigma^2_{\mathrm{H}} + \mathrm{Cov[L,H]})/2\sigma^2_{\mathrm{SMRT}} = 2(\sigma^2_{\mathrm{H}} + \mathrm{Cov[L,H]})/\sigma^2_{\mathrm{L+H}}; \\
&\rho_{\mathrm{FE.SMRT}} = (\sigma^2_{\mathrm{L}} - \sigma^2_{\mathrm{H}})/2\sigma_{\mathrm{FE}}\sigma_{\mathrm{SMRT}} = (\sigma^2_{\mathrm{L}} - \sigma^2_{\mathrm{H}})/\sigma_{\mathrm{L-H}}\sigma_{\mathrm{L+H}}; \text{ and} \\
&b_{\mathrm{FE.SMRT}} = (\sigma^2_{\mathrm{L}} - \sigma^2_{\mathrm{H}})/2\sigma^2_{\mathrm{SMRT}} = 2(\sigma^2_{\mathrm{L}} - \sigma^2_{\mathrm{H}})/\sigma^2_{\mathrm{L+H}} = b_{\mathrm{L.SMRT}} - b_{\mathrm{H.SMRT}}.
\end{aligned}$$

Several properties of the correlations and slopes become apparent from these equations. It will be convenient to assume for the moment that the variance of low-frequency RTs (across subjects) is greater than the variance of high-frequency RTs. Under this assumption it can be seen that

1. The correlation between low-frequency RT and mean RT will never be negative since the covariance can never be larger than the larger variance, which, by assumption, is that of low-frequency RT;

2. The slope of the function relating low-frequency RT to mean RT will always be larger than the slope relating high-frequency RT to mean RT;

3. The slope of the function relating the frequency effect to mean RT is the difference between the slopes for low- and high-frequency RTs;

4. The magnitude of the difference in slopes is a function of the difference between the variances of low- and high-frequency RTs; and

5. As noted in the text, the numerator of the slope of the function relating the frequency effect to mean RT is the difference between the variances of low- and high-frequency RTs.

(Manuscript received August 2, 1995;
revision accepted for publication May 4, 1997.