Journal of Experimental Psychology:
Human Perception and Performance
1996, Vol. 22, No. 6, 1331-1356

0096-1523/96/$3.00

Effects of Polysemy in Lexical Decision and Naming: An Alternative to Lexical Access Accounts

Yasushi Hino
Chukyo University

Stephen J. Lupker
University of Western Ontario

The effects of polysemy (number of meanings) and word frequency were examined in lexical decision and naming tasks. Polysemy effects were observed in both tasks. In the lexical decision task, high- and low-frequency words produced identical polysemy effects. In the naming task, however, polysemy interacted with frequency, with polysemy effects being limited to low-frequency words. When degraded stimuli were used in both tasks, the interaction appeared not only in naming but also in lexical decision. Because stimulus degradation also produced an effect of spelling-sound regularity in the lexical decision task, the different relationships between polysemy and frequency appear to be due to whether responding was based primarily on orthographic or phonological codes. As such, the effects of polysemy seem to be due to the nature of task-specific processes. An explanation in terms of M. S. Seidenberg and J. L. McClelland's (1989) and D. C. Plaut and J. L. McClelland's (1993) parallel distributed processing models is proposed.

One of the most fundamental issues in reading research is how a word's meaning is derived from the processing of a visual input. Chumbley and Balota (1984) suggested that essentially all major models of word recognition, such as Morton's (1969) logogen model, Becker's (1980) verification model, and Forster's (1976) lexical search model, assume at least two processes are involved. The first is the process of accessing the lexicon and the second is the process of meaning determination. The verification model and the lexical search model assume that lexical access involves a sequential matching process between information extracted from the visual stimulus and lexical representations, with representations for higher frequency words checked first. The logogen model assumes differential threshold values for the lexical representations depending on word frequency. When the activation of a logogen reaches its threshold, lexical access is accomplished. More important, these models assume that semantic information becomes available only after lexical access. Thus, in isolated word recognition tasks, these models suggest that semantic variables should have little effect on the lexical access process.

The question of whether semantic variables do affect the lexical access process is the general issue being investigated in the present article. First, we wish to be clear about what we mean by lexical access. As suggested by Fodor (1983), there must be some process that initially maps external information that is given as physical signals onto mental representations so that cognitive processes can operate further on that information. In general, models of the sort mentioned above have assumed that this process is the process of mapping visual signals onto lexical (i.e., word-level) representations and that this is a process that is common to most word recognition tasks (e.g., Balota & Chumbley, 1984, 1985; Chumbley & Balota, 1984). These assumptions will form our working hypothesis as well.

The question of whether semantic variables affect lexical access is one that has received considerable attention in the literature. As pointed out by Balota, Ferraro, and Connor (1991), there is now substantial evidence that semantic factors do affect performance in isolated word recognition tasks. On the basis of this evidence, these authors have concluded that lexical access is affected by semantic variables and have explained these effects in terms of the interactive-activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982).

The interactive-activation model assumes that lexical access is accomplished when a word-level unit is activated over a threshold. Each unit is assumed to have its own

---

Yasushi Hino, Department of Psychology, Chukyo University, Nagoya, Aichi, Japan; Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario, Canada.

This research was based on Yasushi Hino's doctoral dissertation at the University of Western Ontario, under the supervision of Stephen J. Lupker. A major portion of this research was presented at the 3rd Annual Meeting of the Canadian Society for Behavior, Brain and Cognitive Science, Toronto, Ontario, Canada, July 17, 1993.

We thank Patrick Brown, Chris Herdman, Albert Katz, Zenon Pylyshyn, and Chris Sears for their comments and Kaname Hosoya, Penny Pexman, Carolynn Racicot, and Lisa Talvack for their assistance in the data collection. We also thank Ken Paap, Greg Simpson, and Pedro Cabral for their comments on earlier versions of the article.

Correspondence concerning this article should be addressed to Yasushi Hino, Department of Psychology, Chukyo University, 101-2 Yagotohonmachi, Syowaku, Nagoya, Aichi 466, Japan. Electronic mail may be sent via the Internet to hino@sccs.chukyo-u.ac.jp.

resting activation level depending on word frequency. The word-level units are connected to higher, meaning-level units by means of bidirectional links. In this model, the partial activation of word-level units is assumed to send activation signals up to meaning-level units before a word-level unit is activated over its threshold. The activation of meaning-level units, in turn, sends activation signals back down to word-level units. This cascading process facilitates a word-level unit being activated over its threshold. In this way, the interactive-activation model can explain the influence of semantic variables on the lexical access process.[1]

The interactive-activation framework, then, provides a relatively straightforward explanation of how semantic variables could affect lexical access. Nonetheless, the questions still remain: Do semantic variables really affect lexical access and, if so, is the interactive-activation explanation correct? To address these questions, however, one must first grapple with the problem of determining whether effects that appear in certain word recognition tasks truly reflect effects that occur during lexical access.

Most word recognition models were developed primarily on the basis of the results from lexical decision experiments. Originally, the lexical decision task was regarded as the principal task for investigating the lexical access process (e.g., Coltheart, 1978; Paap, McDonald, Schvaneveldt, & Noel, 1987). Some researchers, however, have claimed that lexical decision latency contains a postaccess decision making component and, thus, the task is not a particularly good task for examining the lexical access process (e.g., Balota & Chumbley, 1984; Seidenberg, Waters, Sanders, & Langer, 1984; West & Stanovich, 1982). On the other hand, however, Balota and Chumbley (1985) have also pointed out that latencies in naming tasks (the commonly used alternative task) contain postaccess components as well. Therefore, neither the results from the lexical decision task nor those from the naming task seem to provide a pure measure of lexical access.

One way to deal with the problem of having no one optimal task is to use at least two different tasks and look for parallel effects, an approach suggested by a number of researchers (e.g., Andrews, 1989; Balota & Chumbley, 1984, 1985; Carr, Posner, Pollatsek, & Snyder, 1979; Chumbley & Balota, 1984; Meyer, Schvaneveldt, & Ruddy, 1975; Posner, 1978; West & Stanovich, 1982). Andrews, for example, argued that "evaluations of the patterns of influence of different variables under different task conditions provide a means of specifying the locus of the effects observed" (p. 805). This is also the approach taken in the present article.

The main issue investigated in the present article is the effect on lexical access of a specific semantic variable, polysemy (number of meanings). To examine whether polysemy has any effects on lexical access, both naming and lexical decision tasks were conducted. Polysemy effects have been observed in a number of lexical decision experiments (e.g., Jastrzembski, 1981; Jastrzembski & Stanners, 1975; Kellas, Ferraro, & Simpson, 1988; Millis & Button, 1989; Rubenstein, Garfield, & Millikan, 1970; Rubenstein, Lewis, & Rubenstein, 1971; see also Clark, 1973; Forster & Bednall, 1976; and Gernsbacher, 1984, for criticisms). In these studies, the general result has been that latencies are faster for words with multiple meanings (e.g., *lean, right*) than for words with fewer meanings (e.g., *tent, small*).

In one of the first evaluations of this issue, Rubenstein et al. (1970) collected polysemous and nonpolysemous word groups (hereinafter referred to as *ambiguous* and *unambiguous* word groups) and these groups were further divided in terms of word frequency (high, medium, and low) and concreteness. Number of meanings was measured by asking people to write down the first meaning that came to mind for each word and then counting the number of different meanings produced across all participants. Rubenstein et al. then conducted a lexical decision experiment using these stimuli. Their results showed a significant main effect of polysemy as well as a significant main effect of word frequency. A significant interaction between polysemy and concreteness was also observed, but the interaction between polysemy and frequency was not significant.

On the basis of these results, Rubenstein and his colleagues (Rubenstein et al., 1970, 1971) suggested a lexical access account of polysemy effects. According to their model, when a word is presented, a subset of lexical entries is selected, with high-frequency entries having a higher probability of being selected (the marking process). Next, the selected lexical entries are randomly compared with a visual input (the comparison process). Lexical access is accomplished when a match is found between a lexical entry and the visual input. This model accounts for frequency effects in terms of the marking process. The polysemy effects, however, are explained in terms of the nature of the comparison process. Ambiguous words are assumed to have multiple lexical entries. Because the comparison process randomly selects from among the marked entries, the probability of accessing any one entry for an ambiguous word early in processing is greater than that for a word with only a single entry. Thus, this process should finish sooner for ambiguous words. The model, then, assumes that polysemy and frequency affect the speed of lexical access but in different processing stages. The additive relationship between polysemy and frequency that Rubenstein et al. (1970) obtained was, therefore, nicely accounted for by the model.

Jastrzembski and Stanners (1975) and Jastrzembski (1981) also observed polysemy effects in lexical decision experiments using a slightly different measure of polysemy.

---

[1] For the sake of clarity we will avoid using the terms *prelexical, lexical,* or *postlexical effects* because these terms do not distinguish between process and representation. That is, if an effect is called postlexical, it is unclear whether this means that the process that is affected is postlexical or the representation that is causing the effect is postlexical (or both). The focus here is on the *process* of lexical access, that is, the process of accessing a lexical representation, a process that seems to be central to most models, including the interactive-activation model. Thus, the questions concern what variables affect that process and, in particular, whether representations that are further downstream in the flow of information than the lexical system (i.e., semantic representations) can influence that process.

In these experiments, ambiguous and unambiguous words were selected based on the number of definitions listed for each word in an unabridged dictionary. Like Rubenstein et al. (1970), Jastrzembski orthogonally manipulated polysemy (defined in terms of the number of dictionary definitions) and word frequency in a lexical decision experiment. Unlike Rubenstein et al., however, Jastrzembski observed a significant interaction between the two factors. The size of the polysemy effect was larger for low-frequency words (143 ms) than for high-frequency words (79 ms).

Jastrzembski (1981) explained these results in terms of Morton's (1969) logogen model. He assumed that ambiguous words are represented by separate logogens, with one logogen for each meaning. Because more logogens are activated by ambiguous words, the probability of any one logogen reaching threshold by a given time would be greater for ambiguous words than for unambiguous words. Thus, the polysemy effects were assumed to occur as a result of the horse race among logogens within the frequency-sensitive lexical access mechanism. A natural prediction of this position is the Polysemy × Frequency interaction that Jastrzembski observed.

The point to note here is that these studies do not provide a consistent picture concerning the relationship between polysemy and frequency. Although Rubenstein et al. (1970) observed an additive relationship between polysemy and frequency, Jastrzembski (1981) observed a significant interaction between these factors in lexical decision experiments. There were, needless to say, methodological differences between these studies, differences that may account for the different results. Perhaps most notably, there were differences in how polysemy was defined. Whereas Jastrzembski measured polysemy by counting the number of dictionary definitions, Rubenstein et al. measured polysemy by asking people to write down the first meaning that came to mind for each word.

In fact, both of these techniques have been under attack more recently. Gernsbacher (1984), for example, has suggested that Jastrzembski's (1981) technique greatly overestimates the polysemy of most words. Gernsbacher's informal survey revealed that even well-educated individuals, on average, could report only 3 definitions for the word *fudge,* 2 for *gauge,* and 1 for *cadet,* although these words have 15, 30, and 15 dictionary definitions, respectively. Thus, it is unclear whether Jastrzembski's polysemy manipulation appropriately represented the number of meanings that are really accessed.

On the other hand, Rubenstein et al.'s (1970) measure also may be a bit problematic (Millis & Button, 1989). In their experiments, Millis and Button used three different measures of polysemy and examined which, if any, of those measures produced polysemy effects in a lexical decision task. First, as with Rubenstein et al.'s procedure, they asked people to write down the first meaning of each word and they took the total number of meanings that appeared across participants as a measure of polysemy (first-meaning metric). In addition to this measure, they used two other measures that were derived from the task of asking people to write down all the meanings they could think of for each word. For their second measure, they counted the total number of meanings generated across participants, and the total number was taken as a measure of polysemy (total-meaning metric). Finally, the average number of meanings generated across participants was taken as a third measure of polysemy (average-meaning metric). In their lexical decision experiments, polysemy effects were significant in the analyses treating subjects and items as random factors when the total-meaning metric and the average-meaning metric of polysemy were used. When the first-meaning metric of polysemy was used, however, the polysemy effect was somewhat weaker and significant only in the items' analysis. Thus, Millis and Button's results suggest that the total-meaning and average-meaning metrics might be at least a bit better than Rubenstein et al.'s first-meaning metric for measuring polysemy.

A second difference between Rubenstein et al.'s (1970) technique and Jastrzembski's (1981) was that Jastrzembski did not attempt to control for experiential familiarity. Thus, Gernsbacher (1984) suggested that Jastrzembski's polysemy manipulation might have been confounded with the effect of experiential familiarity. The argument is that although Jastrzembski manipulated word frequency using frequency counts for printed text, controlling printed frequency may not be equivalent to controlling the familiarity of a word in everyday experience. Further, Gernsbacher argued that it seems likely that familiarity correlates with polysemy. That is, the more meanings a given word has, the more likely it is to appear in everyday life. If such is the case, the familiarity for ambiguous words might have been higher than that for unambiguous words in Jastrzembski's materials. Because the difference in lexical decision latencies due to a particular difference in familiarity would be expected to be larger for words in the low-frequency range than for words in the high-frequency range, it is quite possible that the interaction between polysemy and frequency in Jastrzembski's study may have been caused by a confound between polysemy and experiential familiarity.

In the present research, the effects of polysemy and word frequency were examined in Experiment 1 using a lexical decision task. Given the conflicting results reported by Rubenstein et al. (1970) and Jastrzembski (1981) and the criticisms leveled at the measures of polysemy used by those authors, in the present study a different measure of polysemy was used (the number-of-meanings ratings technique reported by Kellas et al., 1988). In this task, people are asked to decide whether a given item has *no meaning* (0), *one meaning* (1), or *more than one meaning* (2) and an average is calculated for each word. This type of rating would seem to provide a measure similar to the average-meaning metric of polysemy used by Millis and Button (1989), which, they argue, better reflects the number of meanings that people can access than Rubenstein et al.'s measure. Further, experiential familiarity ratings were also collected for our experimental items to eliminate the possibility of a confound between polysemy and experiential familiarity.

Observing a polysemy effect in a lexical decision task, however, would not provide an answer to the question of whether polysemy affects the lexical access process because, as noted, a lexical decision task is assumed to involve postaccess, decision-making processes as well. Recently, polysemy effects have also been reported in other paradigms (e.g., Balota et al., 1991; Fera, Joordens, Balota, Ferraro, & Besner, 1992; Kellas et al., 1988). Most relevant to the present work, Fera et al. reported polysemy effects in a naming task, with naming latencies being faster for ambiguous words than for unambiguous words. Polysemy and word frequency were orthogonally manipulated in their naming experiment, and similar polysemy effects were observed for high- (14 ms) and low-frequency words (12 ms) in a standard naming task, whereas there were no polysemy effects in a delayed naming task. Because the polysemy effects were found only in a standard naming task, based on Balota and Chumbley's (1984, 1985) arguments that naming latencies are sensitive to lexical access effects, these effects reinforce the idea that polysemy does affect the lexical access process. The existence of polysemy effects in naming, however, does not conclusively indicate that the effects are lexical access effects because the naming task also seems to involve postaccess processes.

In the present research, Experiment 2 was a naming task using the same word stimuli that were used in the lexical decision task of Experiment 1. Thus, Experiments 1 and 2 provided a direct comparison between the effects of polysemy (and word frequency) in lexical decision and naming tasks. Following Balota and Chumbley (1984, 1985) and Chumbley and Balota (1984), the lexical decision task was assumed to consist of the lexical access process and the postaccess, decision-making processes. Similarly, the naming task was assumed to consist of the lexical access process and the postaccess, pronunciation-related processes. In this context, if polysemy affects only the common lexical access process, the identical pattern of results (as a function of frequency) would be expected in the two experiments. Failure to obtain identical patterns of results would suggest that polysemy had, at the very least, an additional locus in one or both tasks.

Experiment 1

Method

Participants. Twenty-six undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or corrected-to-normal vision.

Stimuli. Forty ambiguous–unambiguous word pairs were created. The 40 ambiguous words were selected from Nelson, McEvoy, Walling, and Wheeler (1980) and Cramer (1970). The ambiguous words were selected to have "meaning frequencies" as equiprobable as possible. In particular, all dominant meaning frequencies for the ambiguous words were less than .61, whereas all subordinate meaning frequencies for the ambiguous words were greater than .15.

The term *meaning frequency* refers to the frequency with which a particular meaning is attached to a polysemous word. In general one meaning is usually attached more frequently than other meanings. In the present circumstances, the meaning frequency measure was based on information contained in word association norms (Cramer, 1970; Nelson et al., 1980). These word association norms are based on N individuals' first responses to a word. Assume that, for a particular polysemous word, the number of people who responded with associated words that were related to meaning A was a. The meaning frequency for meaning A would be given as meaning frequency (A) = a/N.

For example, in Nelson et al. (1980), 46 people were asked to write down a word that came to mind when the ambiguous word *port* was given. Thirty people gave responses that were classified by Nelson et al. as related to the meaning "harbor" (e.g., *ship, boat*), and six responses were classified as related to the meaning "beverage" (e.g., *wine, brandy*). Thus, the meaning frequency for the dominant meaning, harbor, was .652 (30/46). Similarly, the meaning frequency for the subordinate meaning, beverage, was .130 (6/46).

Twenty of these ambiguous words were classified as being high-frequency words (more than 80 per million) and 20 were classified as being low-frequency words (less than 30 per million) by using the Kucera and Francis (1967) norms. Each ambiguous word was then paired with an unambiguous word that had a similar word frequency and word length. Most unambiguous words were taken from Rubenstein et al. (1970, 1971). Some others were selected from the Kucera and Francis norms based on the experimenter's intuition as to the number of meanings.

After collection of these 40 word pairs, 22 people were asked to rate the experiential familiarity for each word. The 80 words were randomly ordered and listed in a questionnaire. Each word was accompanied by a 7-point scale ranging from *very unfamiliar* (1) to *very familiar* (7). The participants were asked to rate the experiential familiarity by circling the appropriate number on the scale.

Another 20 people were asked to rate the number of meanings for these words. The procedure used for collecting the number-of-meanings rating data was identical to that used by Kellas et al. (1988). The 40 ambiguous and 40 unambiguous words were randomly ordered and listed in a questionnaire together with 40 nonwords. At the right-hand side of each item, a scale from 0 to 2 was printed. The participants were asked to decide whether the item had *no meaning* (0), *one meaning* (1), or *more than one meaning* (2) by circling the appropriate number on the scale.

Finally, based on the rating data, 15 high-frequency ambiguous–unambiguous word pairs and 15 low-frequency pairs were selected, ensuring that the number-of-meanings rating values for ambiguous words were more than 1.5 and the values for unambiguous words were less than 1.4. The experiential familiarity rating values were quite comparable between ambiguous and unambiguous word groups. In addition, the mean positional bigram frequencies (Mayzner & Tresselt, 1965), the orthographic neighborhood sizes (Coltheart, Davelaar, Jonasson, & Besner, 1977), and word lengths were equated as much as possible across word groups. Thus, four word groups were created by crossing two factors, word frequency (high or low) and polysemy (ambiguous or unambiguous).

To ensure that the four word groups had been selected properly, analyses of variance (ANOVAs) were conducted on the relevant measures: word frequency, word length, orthographic neighborhood size, mean bigram frequency, experiential familiarity ratings, and number-of-meanings ratings. For word frequency, the main effect of frequency was significant, $F(1, 56) = 21.96$, $p < .001$, $MSE = 31{,}435.88$, but the main effect of polysemy, $F(1, 56) = .00$, $MSE = 31{,}435.88$, and the interaction between polysemy and frequency, $F(1, 56) = .00$, $MSE = 31{,}435.88$, were not significant.

The same results were obtained for the experiential familiarity ratings: frequency, $F(1, 56) = 63.12$, $p < .001$, $MSE = .99$; polysemy, $F(1, 56) = .03$, $MSE = .99$; Frequency × Polysemy, $F(1, 56) = .03$, $MSE = .99$. No significant effects were detected for word length: frequency, $F(1, 56) = .00$, $MSE = .24$; polysemy, $F(1, 56) = .00$, $MSE = .24$; Frequency × Polysemy, $F(1, 56) = .00$, $MSE = .24$; for orthographic neighborhood size: frequency, $F(1, 56) = .70$, $MSE = 21.55$; polysemy, $F(1, 56) = 2.78$, $MSE = 21.55$; Frequency × Polysemy, $F(1, 56) = .251$, $MSE = 21.55$; or for mean bigram frequency: frequency, $F(1, 56) = .58$, $MSE = 1{,}002.58$; polysemy, $F(1, 56) = .25$, $MSE = 1{,}002.58$; Frequency × Polysemy, $F(1, 56) = .02$, $MSE = 1{,}002.58$. For the number-of-meanings ratings, the main effect of polysemy was significant, $F(1, 56) = 763.03$, $p < .001$, $MSE = .01$, but the main effect of frequency, $F(1, 56) = .18$, $MSE = .01$, and the interaction between polysemy and frequency, $F(1, 56) = 1.65$, $MSE = .01$, were not significant.

The experimental word stimuli are listed in Appendix A. The statistical characteristics of these words are given in Table 1. The experimental word stimuli were all four or five letters long. In addition to the experimental word stimuli, 20 filler word stimuli and 80 nonword stimuli were added. Thus, the entire stimulus set consisted of 160 stimuli. All the nonwords were pronounceable nonwords and were created by replacing one letter from actual words. The mean length of the nonwords was 4.40, ranging from 3 to 7. The mean length of the words (experimental + filler) was 4.45, ranging from 3 to 6.

As noted, the present manipulation of polysemy was based on the number-of-meanings ratings. The above analysis indicates that there was no significant difference between high- and low-frequency ambiguous words (1.79 vs. 1.83) on this measure, nor was there any difference between the high- and low-frequency unambiguous words (1.08 vs. 1.05) on this measure. On the other hand, the differences between the dominant and subordinate meaning frequencies were not identical for high- and low-frequency ambiguous words. The difference in meaning frequency between the dominant and subordinate meanings was .19 for high-frequency ambiguous words and .09 for low-frequency ambiguous words (see Table 1), values that were significantly different, $t(28) = 2.90$, $p < .01$. Because there is a possibility that a larger difference in meaning frequency may imply that the word tends to be less ambiguous, the manipulation of polysemy may have been weaker for high-frequency words than for low-frequency words in spite of the equivalent number-of-meanings ratings. Thus, there is a possibility that any observed interaction between polysemy and frequency could be due to this factor.

To address this issue, the data in Experiments 1 and 2 were also analyzed by excluding four high-frequency word pairs (*watch–event, base–loss, shot–clay,* and *fine–food*) to equate the differences in meaning frequency. By excluding these word pairs, the difference in meaning frequencies decreased to .13 for high-frequency ambiguous words and was then comparable to the difference for low-frequency ambiguous words, $t(24) = 1.64$, $p > .10$. The results from neither Experiment 1 nor Experiment 2, however, changed when these four word pairs were removed, suggesting that the nonequivalent differences in meaning frequency between the 15 high- and the 15 low-frequency ambiguous words are unimportant.

Procedure. Participants were tested individually in a normally lit room. Participants were asked to make a word–nonword discrimination for a stimulus appearing on a video monitor (CMS-3436, Multiscan Monitor) by pressing either the word or nonword key. They were also told that their response should be made as quickly and as accurately as possible. Ten practice trials were given prior to the 160 experimental trials. During the practice trials, participants were informed about their lexical decision latency and whether the response was correct after each trial. No feedback information was given during the experimental trials. The order of stimulus presentation for the experimental trials was randomized for each participant.

Each trial was initiated with a 50-ms 400-Hz beep signal. Following the beep, a fixation point appeared at the center of the video monitor. One second after the onset of the fixation point, a stimulus was presented in capital letters above the fixation point. The fixation point and the stimulus were presented in white color on a black background at a luminance of 12 lux (as measured from a 10 mm × 10 mm square at a 0 cm distance by a United Detector Technology, Inc., UDT-40X Opto-Meter in a darkened room). Participants were seated in front of the video monitor at a distance of about 50 cm and asked to respond to the stimulus by pressing either the word or nonword key on the response box interfaced to a microcomputer (AMI 386 Mark II). The word response was made using the participant's dominant hand. The participant's response terminated the presentation of the stimulus and the fixation point. The lexical decision latencies from the onset of the stimulus to the participant's keypress and whether the response was correct were automatically recorded by the microcomputer. The intertrial interval was 3 s.

Results

When a lexical decision latency was less than 250 ms or greater than 1,400 ms, the trial was considered an error. Thus, 17 data points (0.41%) were considered as errors and excluded from the analyses of lexical decision latencies. Mean lexical decision latencies for correct responses and

Table 1
Statistical Characteristics for the Stimuli in Each Condition

Condition frequency–ambiguity	Mean word frequency	Word length	N	BF	FAM	NOM[a]	DOM	SUB
Low–ambiguous	14.2	4.33	7.53	53.26	2.62	1.79	0.50	0.41
Low–unambiguous	14.5	4.33	6.13	50.37	2.62	1.07		
High–ambiguous	226.67	4.33	9.13	60.71	4.62	1.83	0.49	0.30
High–unambiguous	231.13	4.33	6.53	55.42	4.71	1.05		

Note. N = orthographic neighborhood size; BF = positional bigram frequency; FAM = experiential familiarity rating; NOM = number-of-meanings rating; DOM = dominant meaning frequency; SUB = subordinate meaning frequency.
[a] Mean NOM rating for the 40 nonwords was 0.016.

mean error rates were calculated across subjects and items separately. The mean lexical decision latencies and error rates (based on the 60 experimental word trials) averaged across subjects are presented in Table 2.

Subject and item means of lexical decision latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of lexical decision latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 25) = 101.79$, $p < .001$, $MSE = 1{,}078.86$; $F_i(1, 56) = 37.84$, $p < .001$, $MSE = 1{,}907.96$, reflecting the fact that lexical decision latencies were faster for high-frequency words than for low-frequency words. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 25) = 4.68$, $p < .05$, $MSE = 991.14$, although not in the items' analysis, $F_i(1, 56) = 2.45$, $p > .10$, $MSE = 1{,}907.96$. Thus, lexical decision latencies were faster for ambiguous words than for unambiguous words. The interaction between polysemy and frequency was not significant in either analysis, $F_S(1, 25) = .00$, $p > .10$, $MSE = 660.00$; $F_i(1, 56) = .11$, $p > .10$, $MSE = 1{,}907.96$.[2]

In the analyses of error rates, the main effect of frequency was again significant in both analyses, $F_S(1, 25) = 30.88$, $p < .001$, $MSE = .0036$; $F_i(1, 56) = 10.15$, $p < .01$, $MSE = .0063$, reflecting the fact that responses to high-frequency words were more accurate than responses to low-frequency words. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 25) = 13.33$, $p < .001$, $MSE = .0027$, and marginally significant in the items' analysis, $F_i(1, 56) = 3.28$, $p < .10$, $MSE = .0063$. Further, the interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 25) = 9.25$, $p < .01$, $MSE = .0029$, although not in the items' analysis, $F_i(1, 56) = 2.44$, $p > .10$, $MSE = .0063$, reflecting a tendency for the responses to ambiguous words to be more accurate than responses to unambiguous words in the low-frequency word condition, whereas there was no such tendency in the high-frequency word condition.

Discussion

Polysemy and frequency effects were observed for both lexical decision latencies and error rates. The interaction between polysemy and frequency was not significant for lexical decision latency data although it did reach significance for error rates. The interaction in the error rate data seems to have been caused by one item in the low-frequency unambiguous word condition, *veto*. The error rate for this item was .54. Therefore, excluding this item and its paired ambiguous word (*hail*), ANOVAs were again computed. In the analyses of error rates, the significant interaction between polysemy and frequency disappeared, $F_S(1, 25) = 3.14$, $p > .05$, $MSE = .0026$; $F_i(1, 54) = 1.51$, $p > .10$, $MSE = .0031$, although the pattern of results on lexical decision latencies did not change, frequency: $F_S(1, 25) = 92.64$, $p < .001$, $MSE = 996.94$; $F_i(1, 54) = 37.76$, $p < .001$, $MSE = 1{,}443.22$; polysemy: $F_S(1, 25) = 5.07$, $p < .05$, $MSE = 961.12$; $F_i(1, 54) = 2.43$, $p > .10$, $MSE = 1{,}443.22$; Frequency × Polysemy: $F_S(1, 25) = .01$, $p > .10$, $MSE = 773.76$; $F_i(1, 54) = .03$, $p > .10$, $MSE = 1{,}443.22$.

The lack of an interaction between polysemy and frequency on lexical decision latencies replicates the results of Rubenstein et al. (1970); however, it contrasts with the results from Jastrzembski (1981), who obtained a significantly larger polysemy effect for low-frequency words (143 ms) than for high-frequency words (79 ms). What should be noted is that the present study is an improvement on Jastrzembski's in two important ways. First, as noted, Jastrzembski's study and the present study differed in how polysemy was defined. Gernsbacher (1984) pointed out that because Jastrzembski manipulated polysemy in terms of the number of dictionary definitions, his definition might not appropriately represent the number of meanings that actually can be accessed. The number-of-meanings rating technique (Kellas et al., 1988) used here would not be susceptible to that same criticism. Second, because Jastrzembski did not control experiential familiarity, as also noted by Gernsbacher, the familiarity for his ambiguous words might have been higher than that for his unambiguous words. Assuming a logarithmic relationship between lexical decision latencies and familiarity, the difference in lexical decision latencies due to a particular difference in familiarity should be larger for words in the low-frequency range than for words in the high-frequency range. On the basis of the present results and those of Rubenstein et al., it appears that the interaction between polysemy and frequency in Jastrzembski's study was caused either by this confound between polysemy and experiential familiarity or by his problematic definition of polysemy (or both).

Table 2
Mean Lexical Decision Latencies (in Milliseconds) and Error Rates in Experiment 1

	Word frequency				
	Low		High		
Polysemy	*M*	Error rate	*M*	Error rate	Difference
Ambiguous	613	.054	548	.021	+65
Unambiguous	626	.123	561	.026	+65
Difference	+13		+13		

Note. Mean lexical decision latency and error rate for nonwords were 686 ms and .060, respectively.

Experiment 2

Method

Participants. Twenty-six undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or

---

[2] As will be discussed later, although results for items' analyses will be reported for the interested reader, the subjects' analyses are the more appropriate analyses in the present circumstances. Thus, all conclusions will be based on the results from subjects' analyses.

corrected-to-normal vision. None had participated in Experiment 1.

Stimuli. The stimuli were the 80 word stimuli used in Experiment 1.

Procedure. Participants were tested individually. Participants were asked to name a word aloud, which appeared on a video monitor, as quickly and as accurately as possible. Ten practice trials were given prior to the 80 experimental trials. During the practice trials, participants were informed of their naming latency after each trial. No feedback was given during the experimental trials. The order of the stimulus presentation for the experimental trials was randomized for each participant.

On each trial, the stimulus was presented in the same manner and at the same luminance as in Experiment 1. Participants' vocal responses were registered by a microphone connected to a voice key interfaced to the microcomputer used in Experiment 1. (The same microcomputer was also used in all the other experiments reported in this article.) A vocal response terminated the stimulus presentation. Latency was measured from the onset of the stimulus to the onset of the response. An experimenter sat behind the participant and recorded errors. The intertrial interval was 3 s.

Results

A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. The mechanical errors were excluded from the data analyses. There were 27 (1.30%) mechanical errors in total. In addition, when a reaction time was less than 250 ms or more than 1,000 ms, the trial was considered an error. Thus, nine additional data points (0.43%) were considered as errors and removed from the analyses of naming latencies. Mean naming latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean naming latencies and error rates (based on the 60 experimental word trials) averaged across subjects are presented in Table 3.

Subject and item means for naming latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of naming latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 25) = 36.35$, $p < .001$, $MSE = 338.81$; $F_i(1, 56) = 9.55$, $p < .01$, $MSE = 781.41$, reflecting the fact that naming latencies were faster for high-frequency words than for low-frequency words. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 25) = 16.45$, $p < .001$, $MSE = 200.79$, although not in the items' analysis, $F_i(1, 56) = 1.96$, $p > .10$, $MSE = 781.41$. Thus, naming latencies were faster for ambiguous words than for unambiguous words. In addition, the interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 25) = 8.27$, $p < .01$, $MSE = 332.19$, although not in the items' analysis, $F_i(1, 56) = 1.76$, $p > .10$, $MSE = 781.41$. Newman–Keuls tests (based on subject means) were used to examine the difference between ambiguous and unambiguous words for each frequency condition. The polysemy effect was significant for low-frequency words, $q(2, 25) = 6.03$, $p < .01$, but not for high-frequency words, $q(2, 25) = .28$.

Table 3
Mean Naming Latencies (in Milliseconds) and Error Rates in Experiment 2

	Word frequency				
	Low		High		
Polysemy	*M*	Error rate	*M*	Error rate	Difference
Ambiguous	469	.036	457	.010	+12
Unambiguous	490	.049	458	.013	+32
Difference	+21		+1		

In the analyses of error rates, the only significant effect was the main effect of frequency in both analyses, $F_S(1, 25) = 18.23, p < .001, MSE = .0014$; $F_i(1, 56) = 7.72, p < .01$, $MSE = .0019$. Thus, the responses to high-frequency words were more accurate than those to low-frequency words. Neither the main effect of polysemy, $F_S(1, 25) = .79$, $p > .10$, $MSE = .0020$; $F_i(1, 56) = .48, p > .10, MSE = .0019$, nor the interaction between polysemy and frequency, $F_S(1, 25) = .28, p > .10, MSE = .0024$; $F_i(1, 56) = .20, p > .10$, $MSE = .0019$, was significant in either analysis.[3]

Discussion

In the naming task of Experiment 2, polysemy not only affected naming latencies but also interacted with frequency. Polysemy effects appeared only for low-frequency words.

The present results in the naming task appear to be inconsistent with those of Fera et al. (1992). In their standard naming experiment, they obtained polysemy effects for high- and low-frequency words. This apparent inconsistency is probably due to their weaker manipulation of word frequency. According to their description of their stimuli, they had 60 high- and 60 low-frequency words, based on a frequency cutoff of 30 per million (Kucera & Francis, 1967). This cutoff value is substantially lower than the 80 per million cutoff used in the present studies. Thus, it is quite likely that Fera et al. observed polysemy effects for high-frequency words because of their weak frequency manipulation.

Although polysemy affected both naming and lexical decision latencies, the important point to note is that the patterns of results were task specific. There are several possibilities for explaining these results based on task component arguments. There is the possibility, for example, that the different patterns of results are due to articulation onset differences between word groups. Because the initial phonemes were not matched between word groups, the differences in initial phonemes may have differentially affected naming latencies. Assuming that the additive relationship between polysemy and frequency observed in the lexical decision task is due to the lexical access process, any articulation differences that may arise in a naming task may

---

[3] In Experiments 2, 4, 6, and 7, the statistical analyses were redone with the word pair (*hail–veto*) removed, as was done in Experiment 1. In all these experiments, the pattern of results in the latency analysis was unchanged by excluding these items.

change the data pattern. In particular, the pattern may change from additive to the interactive one observed in Experiment 2 if initial articulation takes longer for high-frequency ambiguous words than for high-frequency unambiguous words. To examine this hypothesis, Experiment 3 was a delayed naming task using the stimuli from Experiment 2. If the results of Experiment 2 were due to articulation onset differences between word groups, there should be a reverse polysemy effect for the high-frequency words in a delayed naming task.

Experiment 3

Method

Participants. Eighteen undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or corrected-to-normal vision. None had participated in the previous experiments.

Stimuli. The stimuli were the 60 experimental word stimuli used in Experiments 1 and 2. Three delay conditions (700, 1,000, and 1,300 ms) were used in a within-subject manipulation. The 60 stimuli were divided into three stimulus sets. Each set contained 5 (of the 15) words from each of the four word conditions. Each stimulus set was used in a different delay condition. The assignment of stimulus sets to delay conditions was counterbalanced over three groups of participants. There were 6 participants in each group.

Procedure. Participants were tested individually. They were told that they would see words on the screen and that they should name the word as quickly and accurately as possible as soon as it was surrounded by brackets. Twelve practice trials were given prior to the 60 experimental trials. During the practice trials, participants were informed of their naming latency after each trial. No feedback was given during the experimental trials. The order of stimulus presentation for the experimental trials was randomized for each participant.

Each trial was initiated with a 50-ms 400-Hz beep signal. Following the beep, a fixation point appeared at the center of the display. One second after the onset of the fixation point, a stimulus was presented in capital letters above the fixation point. The stimulus was then surrounded by brackets either 700, 1,000, or 1,300 ms after the onset of the stimulus. The participant's vocal response terminated the stimulus presentation, and the naming latency from the onset of the brackets to the onset of the participant's response was recorded. An experimenter sat behind the participant and recorded errors. The intertrial interval was 3 s.

Results

A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. The mechanical errors were excluded from the data analyses. There were seven (0.65%) mechanical errors in total. In addition, when a reaction time was less than 50 ms or more than 1,000 ms, the trial was considered an error. Thus, 17 additional data points (1.57%) were considered as errors and removed from the analyses of naming latencies. Mean naming latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean naming latencies and error rates averaged across subjects are presented in Table 4.

Subject and item means for naming latencies and error rates were submitted to 2 (polysemy) × 2 (frequency) × 3 (delay) ANOVAs separately. In the subjects' analyses, these were within-subject factors. In the items' analyses, polysemy and frequency were between-item factors and delay was a within-item factor.

In the analyses of naming latencies, the main effect of delay was significant both in the subjects' and the items' analyses, $F_S(2, 34) = 25.15$, $p < .001$, $MSE = 1{,}466.58$; $F_i(2, 112) = 32.88$, $p < .001$, $MSE = 978.16$, reflecting the fact that naming latencies decrease with longer delays. The main effect of frequency was also significant both in the subjects' and the items' analyses, $F_S(1, 17) = 10.02$, $p < .01$, $MSE = 1{,}091.25$; $F_i(1, 56) = 5.24$, $p < .05$, $MSE = 1{,}547.24$, reflecting the fact that naming latencies were faster for high-frequency words than for low-frequency words. Further, the interaction between frequency and delay was significant in the subjects' analysis, $F_S(2, 34) = 4.69$, $p < .05$, $MSE = 522.71$, although not in the items' analysis, $F_i(2, 112) = 1.75$, $p > .10$, $MSE = 987.16$. No other effects approached significance (all Fs < 2.20). Newman–Keuls tests (based on subject means) were used to examine the difference between high- and low-frequency words for each delay condition. The frequency effects were significant in the 700-ms delay condition, $q(2, 34) = 4.36$, $p < .01$, and in the 1,000-ms delay condition, $q(2, 34) = 3.22$, $p < .05$, but not in the 1,300-ms delay condition, $q(2, 25) = .39$.

In the analyses of error rates, the main effect of delay was marginally significant in the items' analysis, $F_i(2, 112) = 2.49$, $p < .09$, $MSE = .0049$, although not in the subjects' analysis, $F_S(2, 34) = 2.44$, $p > .10$, $MSE = .0060$. No other effects were significant in either analysis (all Fs < 2.70).

Discussion

The results from the delayed naming task clearly indicate that there is no difference in delayed naming latencies between our ambiguous and unambiguous word groups (either high or low frequency). Thus, the different patterns of results in Experiments 1 and 2—in particular, the lack of

Table 4
Mean Naming Latencies (in Milliseconds) and Error Rates in Experiment 3

	Delay (ms)					
	700		1,000		1,300	
Frequency–polysemy	*M*	Error rate	*M*	Error rate	*M*	Error rate
Low–ambiguous	347	.033	293	.011	298	.044
Low–unambiguous	339	.000	308	.044	286	.033
Difference	−8		+15		−12	
High–ambiguous	316	.011	282	.011	289	.044
High–unambiguous	322	.033	284	.044	292	.067
Difference	+6		+2		+3	

a polysemy effect for the high-frequency words in Experiment 2—cannot be explained in terms of articulation onset differences.

Another possible explanation for the different results in Experiments 1 and 2 would be based on contributions of the postaccess decision-making processes in the lexical decision task. Let us assume that the lexical access process and the decision-making process in the lexical decision task are both influenced by polysemy. As Balota and Chumbley (1984) suggested, if naming latency is a better measure of lexical access, we could argue that the interaction between polysemy and frequency on naming latencies occurred during lexical access. Because lexical decision latencies were assumed to consist of the lexical access and the decision-making processes, the additive pattern of results between polysemy and frequency should be due to contributions from both processes. Therefore, it would be possible to argue that the relationship between polysemy and frequency changed from interactive to additive because the task-specific decision-making component produced a polysemy effect for high-frequency words in the lexical decision task.

An alternative possibility would be that polysemy actually does not affect the lexical access process but rather differentially affects the postaccess processes in each task. That is, assuming that the lexical decision task consists of the lexical access process and a postaccess decision-making process, polysemy may only affect the decision-making process. The naming task can be assumed to consist of four processes: (a) the lexical access process, (b) a process of retrieving phonological representations, (c) a process of translating phonological representations into articulatory programs, and (d) a process of executing the articulatory programs to produce overt pronunciations. Because there was no effect of polysemy in the delayed naming task, the interactive effect between polysemy and frequency may be due to Process b or to Process c. Thus, the essence of this hypothesis is that the postaccess processes specific to each task are responsible for the different results between lexical decision and naming.

These two hypotheses were examined in Experiment 4 using a go–no go naming task. The go–no go naming task is a variation of the lexical decision task in which overt pronunciations are required only for word stimuli. That is, participants are asked to read a stimulus aloud only when the stimulus is a word. According to the first of the above hypotheses, a key variable is whether the task requires the decision-making process. Because the go–no go naming task requires the decision-making process (as well as the lexical access process, of course), it is essentially identical to the lexical decision task except in the way participants respond to the stimuli. As such, polysemy should be additive with frequency, as in the standard lexical decision task.

On the other hand, if the interactive pattern of results on naming latencies were due to the process of retrieving phonology or the process of translating phonological representations into articulatory programs, the go–no go naming task should produce an interaction between polysemy and frequency. That is, assume that pronunciation-related processes (Processes b, c, and d) are preceded by lexical access and decision-making processes in the go–no go naming task. Thus, after lexical access is accomplished, a word–nonword decision would be made. At this point, polysemy is assumed to have affected the decision-making process in the same way as in the lexical decision task. That is, ambiguous words should have an advantage over unambiguous words independent of frequency. When the stimulus is a word, however, the pronunciation-related processes then have to be carried out to produce a response. These processes, as in the naming task, are assumed to produce an interaction between polysemy and frequency. Therefore, the final response latencies for word stimuli should show an interaction between polysemy and frequency. Further, because the decision-making process was part of the processing sequence, there should be polysemy effects for both low- and high-frequency words.[4]

Experiment 4

Method

Participants. Thirty-two undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers who had normal or corrected-to-normal vision. None had participated in the previous experiments.

Stimuli. The stimuli were the same as used in Experiment 1.

Procedure. Participants were tested individually. They were asked to name a stimulus aloud into a microphone only if the stimulus was a word. They were also told that their responses should be as rapid and as accurate as possible. Ten practice trials were given prior to the 160 experimental trials. During the practice trials, participants were informed of their reaction time and whether their response was correct after each trial. No feedback information was given during the experimental trials. The order of stimulus presentation for the experimental trials was randomized for each participant.

The stimuli were presented in the same manner and at the same luminance as in Experiment 1. The stimulus remained on the video monitor either until the participant responded or until 2 s had elapsed. The participants' task was to name the stimulus aloud into a microphone connected to a voice key only if it was a word. The response latency was measured from the onset of the stimulus to the onset of the participants' response. An experimenter sat behind the participant and recorded errors. The intertrial interval was 3 s.

Results

A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some

---

[4] Note that this prediction is based on the assumption that individuals essentially carry out these tasks in a sequential order. That is, their decision-making processes are virtually complete before they initiate their pronunciation-related processes. A different pattern of results would be expected if these processes were executed in parallel. For example, if the two processes were fully overlapping in time so that the pronunciation-related processes finished prior to the completion of the decision-making process, only the additive effects due to the decision-making process should be observed.

extraneous sound triggered the voice key. The mechanical errors were excluded from the data analyses. There were 88 (1.72% of all trials) mechanical errors in total. In addition, for word trials, when a response latency was less than 250 ms or greater than 1,600 ms, the trial was considered an error and excluded from the analyses of response latencies. Thus, 11 additional data points (0.43% of word trials) were considered as errors and removed from the analyses of response latencies. Mean response latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean response latencies and error rates (based on the 60 experimental word trials) averaged across subjects are presented in Table 5.

Subject and item means of response latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analysis of response latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 31) = 97.64$, $p < .001$, $MSE = 2{,}778.32$; $F_i(1, 56) = 39.08$, $p < .001$, $MSE = 3{,}408.75$, reflecting the fact that response latencies were faster for high-frequency words than for low-frequency words. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 31) = 27.23$, $p < .001$, $MSE = 968.73$, and marginally significant in the items' analysis, $F_i(1, 56) = 3.90$, $p < .06$, $MSE = 3{,}408.75$. Thus, response latencies were faster for ambiguous words than for unambiguous words. Further, the interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 31) = 9.04$, $p < .01$, $MSE = 579.17$, although not in the items' analysis, $F_i(1, 56) = .77$, $p > .10$, $MSE = 3{,}408.75$. Newman–Keuls tests (based on subject means) were used to examine the difference between ambiguous and unambiguous words for each frequency condition. The polysemy effect was significant not only for low-frequency words, $q(2, 31) = 9.76$, $p < .01$, but also for high-frequency words, $q(2, 31) = 3.74$, $p < .05$.

In the analyses of error rates, the main effect of frequency was significant in both analyses, $F_S(1, 31) = 28.36$, $p < .001$, $MSE = .0014$; $F_i(1, 56) = 8.15$, $p < .01$, $MSE = .0023$, reflecting the fact that responses for high-frequency words were more accurate than those for low-frequency words. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 31) = 7.98$, $p < .01$, $MSE = .0013$, and nonsignificant in the items' analysis, $F_i(1, 56) = 2.21$, $p > .10$, $MSE = .0023$. The interaction between polysemy and frequency was also significant in the subjects' analysis, $F_S(1, 31) = 5.42$, $p < .05$, $MSE = .0019$, and nonsignificant in the items' analysis, $F_i(1, 56) = 2.17$, $p > .10$, $MSE = .0023$.

Table 5
Mean Response Latencies (in Milliseconds) and Error Rates in Experiment 4

	Word frequency				
	Low		High		
Polysemy	*M*	Error rate	*M*	Error rate	Difference
Ambiguous	658	.024	579	.006	+79
Unambiguous	700	.059	595	.006	+105
Difference	+42		+16		

Note. Mean error rate for nonwords was .030.

Discussion

Polysemy effects appeared for both low- and high-frequency words in the go–no go naming task. In addition, a significant interaction between polysemy and frequency was observed. That is, the size of the polysemy effect for low-frequency words was larger than that for high-frequency words. These results follow exactly the predictions derived from the hypothesis that polysemy effects arise during postaccess, task-specific processes in both lexical decision and naming tasks. Thus, these results provide considerable support for that hypothesis.

With respect to the postaccess processes in naming, because there was no effect of polysemy in the delayed naming task (i.e., Experiment 3), the polysemy effect in the standard naming task is apparently not due to the articulation process. Thus, the interaction between polysemy and frequency in this task seems to be due to either the process of retrieving phonological representations or the process of translating phonological representations into articulatory programs.

Interestingly, the sizes of the polysemy effects obtained in the present experiment were approximately equal to (but slightly larger than) the sum of the effects appearing in the lexical decision and naming tasks in Experiments 1 and 2. For low-frequency words, a 42-ms polysemy effect was observed in the present experiment, which is just a bit larger than the sum of the 13-ms effect in lexical decision and the 21-ms effect in naming. Similarly, for high-frequency words, the 16-ms polysemy effect in the present experiment was close to the 13-ms effect in lexical decision plus the 1-ms effect in naming. If some portion of polysemy effects were localized at the lexical access process, the sum of the polysemy effects between lexical decision and naming should be larger than the effects in the go–no go naming task because adding the effects in lexical decision to those in naming should add the contribution of the lexical access component twice. Thus, this analysis also supports the conclusion that the polysemy effects that appeared in the lexical decision task arose during the postaccess, decision-making processes, whereas the polysemy effects that appeared in the naming task arose during the postaccess, pronunciation-related processes. These results and this account are, of course, inconsistent with previous lexical access accounts of polysemy effects (Balota et al., 1991; Jastrzembski, 1981; Rubenstein et al., 1970, 1971).

By this same type of logic, the results from these experiments also seem to argue against a lexical access account of frequency effects. That is, the sizes of frequency effects in the present experiment were also approximately equal to (and again, slightly larger than) the sum of the frequency effects observed in lexical decision and in naming (for ambiguous words: $79 \text{ ms} \doteq 65 \text{ ms} + 12 \text{ ms}$; for unambiguous words: $105 \text{ ms} \doteq 65 \text{ ms} + 32 \text{ ms}$). These results seem to suggest that frequency effects are also only minimal during lexical access. Rather, frequency effects also appear to be mostly due to the postaccess, task-specific processes.

Given the conclusion that the polysemy effects as well as frequency effects are mostly due to the postaccess processes specific to each task, the next step would be to propose mechanisms for these effects by considering the essential differences between the task-specific processes that are responsible for the different empirical relationships between polysemy and frequency.

It has been argued that different word recognition tasks require different types of representations to accomplish those tasks (e.g., Carr & Pollatsek, 1985; Carr, Pollatsek, & Posner, 1981; Carr et al., 1979; Seidenberg, 1985, 1989; Seidenberg & McClelland, 1989; Waters & Seidenberg, 1985). In particular, Seidenberg and his colleagues have argued that because a naming task requires production of the correct pronunciations, the task explicitly requires people to retrieve a phonological representation. On the other hand, a lexical decision task does not require people to retrieve phonological representations because the task does not require overt pronunciations. Rather, Seidenberg and his colleagues argued that lexical decisions would usually be made based primarily on the orthographic representations of stimuli if the orthographic information provides enough of a clue to discriminate words from nonwords.

The empirical support for this view has been provided by studies concerning the effects of regularity of spelling–sound correspondences. A number of studies (e.g., Andrews, 1982; Baron & Strawson, 1976; Brown, Lupker, & Colombo, 1994; Coltheart, Besner, Jonasson, & Davelaar, 1979; Glushko, 1979; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Stanovich & Bauer, 1978; Waters & Seidenberg, 1985) have shown that naming latencies for exception words (e.g., *have, give*) were longer than those for regular words (e.g., *gave, save, five, dive*). Further, in most of these studies, this regularity effect was limited to low-frequency words. On the other hand, Waters and Seidenberg found that a regularity effect tends not to appear in the lexical decision task unless the stimulus set contains "strange" words, words that have uncommon or unfamiliar spelling patterns (e.g., *aisle, sign, gauge*).

Waters and Seidenberg (1985) argued that including strange words in the stimulus set made orthographically based word–nonword discriminations difficult because including those words increased the overlap between word and nonword distributions on an orthographic familiarity dimension. Because of the difficulty of the orthographically based discriminations, participants had to make greater use of phonological information in making their word–nonword discriminations. Thus, regularity effects appeared because the decision processes were now carried out based primarily on phonological representations.

If one accepts the argument that lexical decisions are usually made based primarily on orthographic information if the orthographic properties of the stimuli provide enough of a clue to discriminate words from nonwords, whereas naming always necessitates the retrieval of phonological information because overt pronunciations are required, it would be possible to argue that the differential patterns of results between lexical decision and naming in the present studies may be due to the type of representations used in the task-specific processes. That is, the use of (predominantly) orthographic information in the lexical decision task may be the factor that produces the additive relationship between polysemy and frequency. On the other hand, it may be the retrieval and use of phonological information in the naming task that produces the Polysemy × Frequency interaction. To investigate this issue, it is first necessary to establish a manipulation that induces participants to shift the representations used in a lexical decision task from primarily orthographically based ones to primarily phonologically based ones.

According to Seidenberg and his colleagues' (e.g., Waters & Seidenberg, 1985) argument about the type of representations used in the lexical decision task, participants use phonological information much more extensively as the basis for lexical decisions when words and nonwords are fairly similar in terms of orthographic familiarity. If such is the case, it might be possible to bias participants to use phonological representations in a lexical decision task by using degraded stimuli. Because stimulus degradation seems to reduce the availability of visual information, orthographic familiarity differences between words and nonwords may decrease and word–nonword discriminations would be somewhat more difficult to make solely on the basis of the orthographic familiarity.

The reader should note that two points are actually being suggested here. The first is that stimulus degradation makes it more difficult to make word–nonword discriminations solely on the basis of orthographic familiarity. The second is that as the result of stimulus degradation, participants will be induced to make more extensive use of phonological information in the lexical decision task. There is evidence both for and against the first point. For example, the lack of a Word Frequency × Stimulus Quality interaction observed by a number of researchers in lexical decision tasks (e.g., Becker & Killion, 1977; Stanners, Jastrzembski, & Westbrook, 1975; Wilding, 1988, Experiment 1) would suggest that frequency and stimulus quality have independent effects in the sense that they affect different stages in the process. Because frequency and familiarity are quite similar factors, a possible implication is that the same is true for familiarity and stimulus quality. On the other hand, Norris (1984) observed an interaction between frequency and stimulus quality and suggested that the failure of others to observe the interaction stems mainly from the weak stimulus quality manipulations they used.

The more important point, however, is the second one—that is, the issue of whether degradation will induce more reliance on phonological information in a lexical decision task. This is the empirical question addressed in Experiments 5A, 5B, and 5C. In particular, the effects of regularity of spelling–sound correspondences were examined in naming and lexical decision tasks with clear stimuli with the expectation that we would replicate Waters and Seidenberg's (1985) results (Experiments 5A and 5B, respectively). Successful replications would involve finding a regularity effect for low-frequency words in the naming task but not in the lexical decision task. Finally, a lexical decision task was conducted in which the same stimuli were presented at a degraded luminance (Experiment 5C). If degrading stimuli in this way induces a change to a more phonologically based strategy, we should also observe a regularity effect for low-frequency words in Experiment 5C.

Experiments 5A and 5B

Method

Participants. Fifty-eight undergraduate students from the University of Western Ontario participated in these experiments for course credit. Twenty-six participants took part in Experiment 5A (naming task) and 32 took part in Experiment 5B (lexical decision task). All were native English speakers and had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. All the stimuli were four letters long. Twenty regular and 20 exception words were selected. One half of the regular and one half of the exception words were high-frequency words (frequency greater than or equal to 70 per million) according to the Kucera and Francis (1967) norms. The other half were low-frequency words (frequency less than or equal to 30 per million). Thus, four word groups were created by crossing two factors, word frequency (high or low) and regularity (regular or exception). For the purpose of excluding strange words from the stimulus set, only words with many orthographic neighbors (more than six neighbors) were used. The orthographic neighborhood size and mean positional bigram frequency were closely equated across word groups. The experiential familiarity rating values were also closely equated between regular and exception word groups. Twenty-four people were asked to rate the experiential familiarity of these words using a 7-point scale ranging from *very unfamiliar* (1) to *very familiar* (7).

To ensure that the four word groups had been selected appropriately, ANOVAs were conducted on the relevant factors: word frequency, orthographic neighborhood size, mean bigram frequency, and experiential familiarity rating. For word frequency, the main effect of frequency was significant, *F*(1, 36) = 49.06, *p* < .001, *MSE* = 16,659.77, but the main effect of regularity, *F*(1, 36) = .00, *MSE* = 16,659.77, and the interaction between regularity and frequency, *F*(1, 36) = .00, *MSE* = 16,659.77, were not significant. The same results were obtained for the experiential familiarity ratings, frequency: *F*(1, 36) = 169.54, *p* < .001, *MSE* = .48; regularity: *F*(1, 36) = .27, *MSE* = .48; Frequency × Regularity: *F*(1, 36) = .20, *MSE* = .48. No significant effects were detected either for orthographic neighborhood size, frequency: *F*(1, 36) = .25, *MSE* = 8.25; regularity: *F*(1, 36) = .03, *MSE* = 8.25; Frequency × Regularity: *F*(1, 36) = .08, *MSE* = 8.25, or for mean bigram frequency, frequency: *F*(1, 36) = .85, *MSE* = 647.15; regularity: *F*(1, 36) = .00, *MSE* = 647.15; Frequency × Regularity: *F*(1, 36) = .87, *MSE* = 647.15.

The experimental word stimuli are listed in Appendix B. The statistical characteristics of these words are given in Table 6. In Experiment 5A, eight filler word stimuli were used in addition to the 40 experimental word stimuli. Thus, the entire stimulus set consisted of 48 word stimuli. The eight filler word stimuli were also words with many orthographic neighbors (more than five neighbors). In addition to these 48 word stimuli, 48 nonword stimuli were used in Experiment 5B. All the nonwords were pronounceable nonwords and were created by replacing one letter from actual words.

Table 6
Statistical Characteristics for the Stimuli in Each Condition in Experiments 5A, 5B, and 5C

Condition frequency–regularity	Mean word frequency	Word length	N	BF	FAM
Low–regular	8.6	4.0	10.7	78.53	2.97
Low–exception	9.3	4.0	10.3	70.63	2.99
High–regular	293.1	4.0	10.9	78.43	5.72
High–exception	296.6	4.0	11.0	85.53	5.93

Note. N = orthographic neighborhood size; BF = positional bigram frequency; FAM = experiential familiarity rating.

Procedure. Participants were tested individually in a normally lit room. All the stimuli and the fixation point were presented in white on a black background at a luminance of 12 lux. The procedure of Experiment 5A (naming task) was identical to that of Experiment 2, and the procedure of Experiment 5B (lexical decision task) was identical to that of Experiment 1.

Results

Experiment 5A (naming task). A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. The mechanical errors were excluded from the data analyses. There were nine (0.72%) mechanical errors in total. In addition, when a reaction time was less than 250 ms or more than 1,000 ms, the trial was considered an error. Thus, three additional data points (0.24%) were considered as errors and removed from the analyses of naming latencies. Mean naming latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean naming latencies and error rates (based on the 40 experimental word trials) averaged across subjects are presented in Table 7.

Subject and item means for naming latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of naming latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 25) = 46.93$, *p* < .001, *MSE* = 589.10; $F_i(1, 36) = 16.28$, *p* < .001, *MSE* = 779.87, reflecting the fact that naming latencies were faster for high-frequency words than for low-frequency words. The main effect of regularity was significant in the subjects' analysis, $F_S(1, 25) = 20.72$, *p* < .001, *MSE* = 261.91, and marginally significant in the items' analysis, $F_i(1, 36) =$

Table 7
Mean Naming Latencies (in Milliseconds) and Error Rates in Experiment 5A

	Word frequency				
	Low		High		
Regularity	*M*	Error rate	*M*	Error rate	Difference
Regular	469	.012	450	.023	+19
Exception	498	.129	451	.008	+47
Difference	+29		+1		

3.93, $p < .10$, $MSE = 779.87$. Thus, naming latencies were faster for regular words than for exception words. In addition, the interaction between regularity and frequency was significant in the subjects' analysis, $F_S(1, 25) = 12.18$, $p < .01$, $MSE = 412.74$, and marginally significant in the items' analysis, $F_i(1, 36) = 3.65$, $p < .10$, $MSE = 779.87$. Newman–Keuls tests (based on subject means) were used to examine the difference between regular and exception words for each frequency condition. The regularity effect was significant for low-frequency words, $q(2, 25) = 7.12$, $p < .01$, but not for high-frequency words, $q(2, 25) = .14$.

In the analyses of error rates, the main effect of frequency was significant in the subjects' analysis, $F_S(1, 25) = 20.00$, $p < .001$, $MSE = .0039$, although not in the items' analysis, $F_i(1, 36) = 1.92$, $p > .10$, $MSE = .0153$. The main effect of regularity was also significant in the subjects' analysis, $F_S(1, 25) = 29.00$, $p < .001$, $MSE = .0023$, although not in the items' analysis, $F_i(1, 36) = 1.64$, $p > .10$, $MSE = .0153$. Further, the interaction between regularity and frequency was significant in the subjects' analysis, $F_S(1, 25) = 46.59$, $p < .001$, $MSE = .0024$, although not in the items' analysis, $F_i(1, 36) = 2.80$, $p > .10$, $MSE = .0153$. The significant interaction reflects the comparatively higher error rates for low-frequency exception words.

Experiment 5B (lexical decision task). When a lexical decision latency was less than 250 ms or greater than 1,400 ms, the trial was considered an error. Thus, four data points (0.13%) were considered as errors and excluded from the analyses of lexical decision latencies. Mean lexical decision latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean lexical decision latencies and error rates (based on the 40 experimental word trials) averaged across subjects are presented in Table 8.

Subject and item means of lexical decision latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of lexical decision latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 31) = 138.13$, $p < .001$, $MSE = 1{,}497.23$; $F_i(1, 36) = 41.55$, $p < .001$, $MSE = 1{,}747.73$, reflecting the fact that lexical decision latencies were faster for high-frequency words than for low-frequency words. Neither the main effect of regularity, $F_S(1, 31) = .01$, $MSE = 1{,}104.76$; $F_i(1, 36) = .03$, $MSE = 1{,}747.73$, nor the interaction between regularity and frequency, $F_S(1, 31) = 1.58$, $MSE = 1{,}112.15$; $F_i(1, 36) = .23$, $MSE = 1{,}747.73$, was significant in either analysis.

Table 8
Mean Lexical Decision Latencies (in Milliseconds) and Error Rates in Experiment 5B

	Word frequency				
	Low		High		
Regularity	*M*	Error rate	*M*	Error rate	Difference
Regular	573	.116	500	.031	+73
Exception	579	.131	492	.013	+87
Difference	+6		−8		

Note. Mean lexical decision latency and error rate for nonwords were 636 ms and .100, respectively.

In the analyses of error rates, the main effect of frequency was again significant in both analyses, $F_S(1, 31) = 56.09$, $p < .001$, $MSE = .0059$; $F_i(1, 36) = 14.74$, $p < .001$, $MSE = .0068$, reflecting the fact that responses to high-frequency words were more accurate than responses to low-frequency words. Neither the main effect of regularity, $F_S(1, 31) = .03$, $MSE = .0027$; $F_i(1, 36) = .00$, $MSE = .0068$, nor the interaction between regularity and frequency, $F_S(1, 31) = 2.20$, $p > .10$, $MSE = .0043$; $F_i(1, 36) = .23$, $MSE = .0068$, was significant in either analysis.

Discussion

The interaction between regularity and frequency appeared in the naming task. In the lexical decision task, neither the main effect of regularity nor the interaction between regularity and frequency was significant. Thus, as argued by Waters and Seidenberg (1985), the results seem to suggest that whereas naming necessitates the retrieval of phonological information because overt pronunciations are required, in general, lexical decisions are based primarily on orthographic information.

Experiment 5C

Experiment 5C was identical to Experiment 5B except that the stimuli were degraded. The question is whether stimulus degradation biases participants to use phonological information in the decision-making process, hence producing a Frequency × Regularity interaction in a lexical decision task.

Method

Participants. Thirty-two undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. The stimuli were the same as used in Experiment 5B. In this experiment, however, all stimuli were presented in a degraded intensity on a video monitor. The degradation was done by reducing the voltage on the red, green, and blue signals of an analog video monitor through digital to analog converter (DAC) register programming (Kliewer, 1988). The luminance of the stimuli was measured in a darkened room from a 10 mm × 10 mm square at a 0 cm distance by a United Detector Technology, Inc., UDT-40X Opto-Meter. All the stimuli were presented at a luminance of 0.036 lux just above a fixation point. The fixation point was located at the center of the video monitor at a luminance of 0.10 lux.

Procedure. Participants were tested individually in a darkened room. The procedure was identical to that of Experiment 5B except that the luminance of the stimuli was reduced.

Results

When a lexical decision latency was less than 250 ms or greater than 2,000 ms,[5] the trial was considered an error. Thus, seven data points (0.23%) were considered as errors and excluded from the analyses of lexical decision latencies. Mean lexical decision latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean lexical decision latencies and error rates (based on the 40 experimental word trials) averaged across subjects are presented in Table 9.

Subject and item means of lexical decision latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of lexical decision latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 31) = 72.53, p < .001, MSE = 4{,}597.59$; $F_i(1, 36) = 63.96, p < .001, MSE = 1{,}643.70$, reflecting the fact that lexical decision latencies were faster for high-frequency words than for low-frequency words. The main effect of regularity was marginally significant in the subjects' analysis, $F_S(1, 31) = 3.78, p < .07, MSE = 1{,}435.07$, although not in the items' analysis, $F_i(1, 36) = .94, p > .10, MSE = 1{,}643.70$. The interaction between regularity and frequency was significant in the subjects' analysis, $F_S(1, 31) = 5.20, p < .05, MSE = 1{,}588.53$, although not in the items' analysis, $F_i(1, 36) = 1.45, p > .10, MSE = 1{,}643.70$. Newman–Keuls tests (based on subject means) were used to examine the difference between regular and exception words for each frequency condition. The regularity effect was significant for low-frequency words, $q(2, 31) = 4.13, p < .01$, but not for high-frequency words, $q(2, 31) = .43$. Thus, the regularity effect was limited to low-frequency words.

In the analyses of error rates, the main effect of frequency was significant in both analyses, $F_S(1, 31) = 30.32, p < .001, MSE = .0064$; $F_i(1, 36) = 13.94, p < .001, MSE = .0044$, reflecting the fact that responses to high-frequency words were more accurate than responses to low-frequency words. Neither the main effect of regularity, $F_S(1, 31) = 2.01, MSE = .0056$; $F_i(1, 36) = .80, MSE = .0044$, nor the interaction between regularity and frequency, $F_S(1, 31) = .06, MSE = .0053$; $F_i(1, 36) = .02, MSE = .0044$, was significant in either analysis.

Table 9
Mean Lexical Decision Latencies (in Milliseconds) and Error Rates in Experiment 5C

	Word frequency				
	Low		High		
Regularity	*M*	Error rate	*M*	Error rate	Difference
Regular	735	.109	649	.028	+86
Exception	764	.088	646	.013	+118
Difference	+29		−3		

Note. Mean lexical decision latency and error rate for nonwords were 830 ms and .090, respectively.

Discussion

When degraded stimuli were used in a lexical decision task, an interaction between regularity and frequency was observed. Thus, the results suggest that when stimuli are degraded, phonological information plays a much more important role in making lexical decisions. Our suggestion is that a strong stimulus degradation, like that used in Experiment 5C, reduces the quality of the orthographic information, making the word stimuli look less wordlike. The result is that orthographic familiarity differences between words and nonwords decrease. As a result, orthographically based lexical decisions become much more difficult, biasing participants to make greater use of phonological information in the decision-making process.

On the basis of these results and this analysis, the stimulus quality manipulation should be useful in attempting to determine why there are different relationships between polysemy and frequency in the lexical decision and naming tasks. In particular, because stimulus degradation seems to bias participants to use phonological information as the basis for making lexical decisions, if the interaction between polysemy and frequency in the naming task is due to the use of phonological representations, the interaction should appear in a lexical decision task with degraded stimuli. To investigate this hypothesis, Experiment 6 was a lexical decision task using degraded stimuli.

Experiment 6

Method

Participants. Thirty undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. The stimuli were the same as used in Experiment 1. These stimuli were presented in a degraded intensity on a video monitor as in Experiment 5C. Thus, all the stimuli were presented at a luminance of 0.036 lux. The luminance of the fixation point was 0.10 lux.

Procedure. Participants were tested individually in a darkened room. The procedure was identical to that in Experiment 5C.

Results

A trial was considered an error if the lexical decision latency was less than 250 ms or more than 2,000 ms. Because 4 participants showed too many errors (more than

---

[5] Different cutoffs were used with clear versus degraded stimuli because reaction times were generally longer when degraded stimuli were used. In the experiments using degraded stimuli (Experiments 5C, 6, and 7), however, the statistical analyses were also conducted using the same cutoffs as in corresponding clear stimulus experiments (Experiments 5B, 1, and 2, respectively). In all these experiments, the pattern of results was unchanged by using these lower cutoffs in spite of the fact that many more reaction times had to be discarded.

15%), their data were excluded from the data analyses. Thus, the data from 26 participants were submitted to the analyses. For those 26, five data points (0.12%) were out of the allowable range mentioned above. Thus, these were regarded as errors and excluded from the analyses of lexical decision latencies. Mean lexical decision latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean lexical decision latencies and error rates (based on the 60 experimental word trials) averaged across subjects are presented in Table 10.

Subject and item means of lexical decision latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of lexical decision latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 25) = 50.58, p < .001, MSE = 3{,}422.09$; $F_i(1, 56) = 27.08, p < .001, MSE = 4{,}164.09$, reflecting the fact that lexical decision latencies were faster to high-frequency words than to low-frequency words. The main effect of polysemy was not significant in either analysis, $F_S(1, 25) = 2.42, p > .10, MSE = 2{,}539.80$; $F_i(1, 56) = 1.01, p > .10, MSE = 4{,}164.09$. The interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 25) = 6.58, p < .025, MSE = 1{,}249.27$, although not in the items' analysis, $F_i(1, 56) = 1.17, p > .10, MSE = 4{,}164.09$. Newman–Keuls tests (based on subject means) were used to examine the difference between ambiguous and unambiguous words for each frequency condition. The polysemy effect was significant for low-frequency words, $q(2, 25) = 4.49, p < .01$, but not for high-frequency words, $q(2, 25) = .35$.

In the analyses of error rates, the main effect of frequency was again significant in both analyses, $F_S(1, 25) = 32.82, p < .001, MSE = .0039$; $F_i(1, 56) = 20.63, p < .001, MSE = .0036$, reflecting the fact that responses to high-frequency words were more accurate than those to low-frequency words. The main effect of polysemy was also significant in both analyses, $F_S(1, 25) = 9.12, p < .01, MSE = .0039$; $F_i(1, 56) = 5.74, p < .025, MSE = .0036$, indicating that responses were more accurate for ambiguous words than for unambiguous words. The interaction between polysemy and frequency was not significant in either analysis, $F_S(1, 25) = 1.86, p > .10, MSE = .0052$; $F_i(1, 56) = 1.54, p > .10, MSE = .0036$.

To examine the effects of stimulus quality directly, combined analyses with the data from Experiment 1 were carried out. The subject and item means of lexical decision latencies and error rates from Experiments 1 and 6 were submitted to 2 (polysemy) × 2 (frequency) × 2 (stimulus quality) ANOVAs separately. In the subjects' analyses, frequency and polysemy were within-subject factors and stimulus quality was a between-subject factor. In the items' analyses, frequency and polysemy were between-item factors and stimulus quality was a within-item factor.

In the analyses of lexical decision latencies, the main effect of stimulus quality was significant in both analyses, $F_S(1, 50) = 87.74, p < .001, MSE = 16{,}692.21$; $F_i(1, 56) = 729.45, p < .001, MSE = 1{,}175.37$, reflecting the fact that lexical decision latencies were slower for degraded stimuli. The main effect of frequency was also significant in both analyses, $F_S(1, 50) = 124.13, p < .001, MSE = 2{,}250.48$; $F_i(1, 56) = 37.31, p < .001, MSE = 4{,}896.68$. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 50) = 6.18, p < .025, MSE = 1{,}765.47$, although not in the items' analysis, $F_i(1, 56) = 1.81, p > .10, MSE = 4{,}896.68$. The two-way interaction between polysemy and frequency was also significant in the subjects' analysis, $F_S(1, 50) = 4.39, p < .05, MSE = 954.64$, although not in the items' analysis, $F_i(1, 56) = .73, p > .10, MSE = 4{,}896.68$. Further, the three-way interaction among polysemy, frequency, and stimulus quality was significant in the subjects' analysis, $F_S(1, 50) = 4.23, p < .05, MSE = 954.64$, although not in the items' analysis, $F_i(1, 56) = 1.30, p > .10, MSE = 1{,}175.37$. Neither the interaction between frequency and stimulus quality, $F_S(1, 50) = 1.59, MSE = 2{,}250.48$; $F_i(1, 56) = 1.91, MSE = 1{,}175.37$, nor the interaction between polysemy and stimulus quality, $F_S(1, 50) = .03, MSE = 1{,}765.47$; $F_i(1, 56) = .01, MSE = 1{,}175.37$, was significant in either analysis.

In the analyses of error rates, the main effect of frequency was significant in both analyses, $F_S(1, 50) = 63.70, p < .001, MSE = .0038$; $F_i(1, 56) = 18.21, p < .001, MSE = .0076$. The main effect of polysemy was also significant in both analyses, $F_S(1, 50) = 21.67, p < .001, MSE = .0033$; $F_i(1, 56) = 5.45, p < .025, MSE = .0076$. The two-way interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 50) = 8.48, p < .01, MSE = .0040$, although not in the items' analysis, $F_i(1, 56) = 2.59, p > .10, MSE = .0076$. All other effects were nonsignificant (all $Fs < 1.00$).

Table 10
Mean Lexical Decision Latencies (in Milliseconds) and Error Rates in Experiment 6

	Word frequency				
	Low		High		
Polysemy	*M*	Error rate	*M*	Error rate	Difference
Ambiguous	779	.064	715	.013	+64
Unambiguous	812	.121	713	.031	+99
Difference	+33		−2		

Note. Mean lexical decision latency and error rate for nonwords were 855 ms and .060, respectively.

Discussion

The major result of Experiment 6 is that the interaction between polysemy and word frequency appeared when the stimuli were degraded. The interaction was due to the fact that the effect of polysemy was limited to low-frequency words. Whereas this interaction is the same one previously observed in a naming task, it stands in contrast to the additive pattern observed in a lexical decision task with clear stimuli. This point is underscored by the significant three-way interaction among polysemy, frequency, and stimulus quality in the combined analysis. That is, this three-way interaction is due to the fact that the interaction between polysemy and frequency appeared when the stimuli were degraded, whereas polysemy was additive with frequency when clear stimuli were used.

The results in the present experiment were actually quite similar to those observed in the naming task of Experiment 2. In both experiments, the polysemy effect was limited to low-frequency words. Further, note that in both experiments the size of the frequency effect for unambiguous words (32 ms in Experiment 2, 99 ms in Experiment 6) was substantially larger than that for ambiguous words (12 ms in Experiment 2, 64 ms in Experiment 6). On the other hand, when clear stimuli were used in lexical decision (Experiment 1), the size of frequency effects was the same for ambiguous and unambiguous words (65 ms). These similarities between the results of the present experiment and those of the naming task provide support for the arguments that stimulus degradation substantially increases participants' use of phonological representations in the decision-making process, and it is the use of phonological representations that gives rise to the interaction of polysemy and frequency.

The three-way interaction among polysemy, frequency, and stimulus quality can also be looked at in a different way. It suggests that the relationship between frequency and stimulus quality is modulated by the polysemy of the words. Most studies investigating the relationship between frequency and stimulus quality in lexical decision have produced additivity (Becker & Killion, 1977; Stanners et al., 1975; Wilding, 1988, Experiment 1). There have, however, also been studies that have shown an interactive relationship between these variables (Norris, 1984; Wilding, 1988, Experiment 2). As noted, Norris has argued that a possible reason for the discrepancy is that researchers who failed to find an interaction did not have a strong degradation manipulation. The present results suggest an alternative, or at least additional, reason for the discrepancy, the polysemy of the words. For ambiguous words, an additive relationship was obtained. The sizes of frequency effects were almost the same for clear (65 ms) and degraded stimuli (64 ms). On the other hand, for unambiguous words, the size of the frequency effect for degraded stimuli (99 ms) was much larger than that for clear stimuli (65 ms). Thus, the possibility exists that researchers reporting no interaction used mainly ambiguous words whereas those reporting an interaction used mainly unambiguous words. To test this hypothesis, of course, it would be necessary to obtain number-of-meanings rating data on the words used in the previous studies. If the hypothesis is correct, the ratings should be lower (i.e., the words should be less polysemous) in the studies showing the interaction.

On the basis of the results from Experiments 1–4, the interaction between polysemy and frequency on naming latencies was assumed to be due to either the process of retrieving phonological representations or the process of translating phonological representations into articulatory programs. As the results of Experiment 6 showed, the same interaction can be obtained in a lexical decision task if the stimuli are degraded. Because the lexical decision task does not require overt pronunciation responses, it is unlikely that the lexical decision task involves a process of translating phonological representations into articulatory programs (even if participants are biased to rely more on phonological representations). Thus, the interaction between polysemy and frequency seems to be due to the process of retrieving phonological representations.

Experiment 7

In Experiment 7, a naming task was conducted with degraded stimuli. This experiment was essentially a manipulation check. In a naming task, the identical interaction between polysemy and frequency should be found regardless of stimulus quality because the naming task does require the retrieval of phonological representations to produce overt pronunciations.

Method

Participants. Thirty undergraduate students from the University of Western Ontario participated in this experiment for course credit. All were native English speakers and had normal or corrected-to-normal vision. None had participated in any of the previous experiments.

Stimuli. The stimuli were the same as used in Experiment 2. As in Experiment 6, all the stimuli were presented at a degraded intensity on a video monitor. The luminance of the stimuli was identical to that of Experiment 6.

Procedure. Participants were tested individually in a darkened room. The procedure was identical to that of Experiments 2 and 5A except that the luminance of the stimuli was reduced.

Results

A trial was considered a mechanical error if the participant's vocal response failed to trigger the voice key or some extraneous sound triggered the voice key. The mechanical errors were excluded from the data analyses. Further, a trial was considered an error if the naming latency was less than 250 ms or more than 1,300 ms. Because 4 participants showed too many errors (more than 15%), their data were excluded from the analyses. Thus, the data from 26 participants were submitted to the analyses. For those 26, there were 30 (1.44%) mechanical errors in total and 15 data points (0.72%) were out of the allowable range mentioned above. Mean naming latencies for correct responses and mean error rates were calculated across subjects and items separately. The mean naming latencies and error rates (based on the 60 experimental word trials) averaged across subjects are presented in Table 11.

Subject and item means of naming latencies and error rates (based on the experimental word trials) were submitted to separate ANOVAs. In the analyses of naming latencies, the main effect of frequency was significant both in the subjects' and the items' analyses, $F_S(1, 25) = 49.89$, $p < .001$, $MSE = 1{,}700.75$; $F_i(1, 56) = 18.11$, $p < .001$, $MSE = 2{,}797.31$, reflecting the fact that naming latencies were faster for high-frequency words than for low-frequency

Table 11
Mean Naming Latencies (in Milliseconds) and Error Rates in Experiment 7

	Word frequency				
	Low		High		
Polysemy	*M*	Error rate	*M*	Error rate	Difference
Ambiguous	663	.051	622	.034	+41
Unambiguous	686	.040	612	.021	+74
Difference	+23		−10		

words. The main effect of polysemy was not significant in either analysis, $F_S(1, 25) = 1.98$, $p > .10$, $MSE = 537.73$; $F_i(1, 56) = .41$, $p > .10$, $MSE = 2{,}797.31$. The interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 25) = 4.38$, $p < .05$, $MSE = 1{,}566.81$, although not in the items' analysis, $F_i(1, 56) = 1.16$, $p > .10$, $MSE = 2{,}797.31$. Newman–Keuls tests (based on subject means) were used to examine the difference between ambiguous and unambiguous words for each frequency condition. The polysemy effect was significant for low-frequency words, $q(2, 25) = 2.92$, $p < .05$, but not for high-frequency words, $q(2, 25) = 1.27$.

In the analyses of error rates, the main effect of frequency was significant in the subjects' analysis, $F_S(1, 25) = 4.44$, $p < .05$, $MSE = .0020$, although not in the items' analysis, $F_i(1, 56) = 1.67$, $p > .10$, $MSE = .0029$. The main effect of polysemy was marginally significant in the subjects' analysis, $F_S(1, 25) = 3.37$, $p < .08$, $MSE = .0011$, and nonsignificant in the items' analysis, $F_i(1, 56) = .79$, $p > .10$, $MSE = .0029$. The interaction between polysemy and frequency was not significant in either analysis, $F_S(1, 25) = .01$, $MSE = .0013$; $F_i(1, 56) = .00$, $MSE = .0029$.

To examine the effects of stimulus quality directly, combined analyses with the data from Experiments 2 and 7 were carried out. The subject and item means of naming latencies and error rates from Experiments 2 and 7 were submitted to 2 (polysemy) × 2 (frequency) × 2 (stimulus quality) ANOVAs separately. In the subjects' analyses, polysemy and frequency were within-subject factors and stimulus quality was a between-subject factor. In the items' analyses, polysemy and frequency were between-item factors and stimulus quality was a within-item factor.

In the analyses of naming latencies, the main effect of stimulus quality was significant in both analyses, $F_S(1, 50) = 64.29$, $p < .001$, $MSE = 25{,}448.90$; $F_i(1, 56) = 789.91$, $p < .001$, $MSE = 1{,}171.19$, reflecting the fact that naming latencies were slower for degraded stimuli. The main effect of frequency was also significant in both analyses, $F_S(1, 50) = 79.28$, $p < .001$, $MSE = 1{,}019.78$; $F_i(1, 56) = 20.14$, $p < .001$, $MSE = 2{,}407.53$. The main effect of polysemy was significant in the subjects' analysis, $F_S(1, 50) = 11.00$, $p < .01$, $MSE = 369.26$, although not in the items' analysis, $F_i(1, 56) = 1.11$, $p > .10$, $MSE = 2{,}407.53$. The interaction between frequency and stimulus quality was significant in both analyses, $F_S(1, 50) = 15.92$, $p < .001$, $MSE = 1{,}019.78$; $F_i(1, 56) = 8.21$, $p < .01$, $MSE = 1{,}171.19$, reflecting the fact that frequency effects were larger when the stimuli were degraded. The interaction between polysemy and frequency was significant in the subjects' analysis, $F_S(1, 50) = 9.64$, $p < .01$, $MSE = 949.50$, although not in the items' analysis, $F_i(1, 56) = 1.83$, $p > .10$, $MSE = 2{,}407.53$. All other effects were nonsignificant (all *F*s < 1.00). The significant interaction between polysemy and frequency and the nonsignificant interaction among polysemy, frequency, and stimulus quality indicated that polysemy effects appeared only for low-frequency words regardless of stimulus quality.

In the analyses of error rates, the main effect of frequency was significant in both analyses, $F_S(1, 50) = 18.91$, $p < .001$, $MSE = .0017$; $F_i(1, 56) = 5.69$, $p < .025$, $MSE = .0032$. The interaction between polysemy and stimulus quality was marginally significant in the subjects' analysis, $F_S(1, 50) = 3.26$, $p < .08$, $MSE = .0016$, although not in the items' analysis, $F_i(1, 56) = 1.90$, $p > .10$, $MSE = .0016$. No other effects were significant (all *F*s < 1.60).

Discussion

The results from the present experiment were quite similar to those from Experiments 2 and 6. Polysemy interacted with frequency, and polysemy effects were limited to low-frequency words. (Note also that, as in Experiment 2, the size of the frequency effect for unambiguous words [74 ms] was again larger than that for ambiguous words [41 ms].) Thus, when the stimuli are degraded, polysemy interacts with frequency regardless of task type, but when clear stimuli are used, the interaction appears only in the naming task, with polysemy being additive with frequency in the lexical decision task. These results clearly indicate that the interactive relationship between polysemy and frequency is neither specific to a particular task nor specific to a particular stimulus quality. Rather, the interaction seems to occur when the task involves the retrieval of phonological representations. Naming tasks require the retrieval of phonological representations regardless of stimulus quality because participants have to produce overt pronunciations. In a lexical decision task, however, phonological representations were used extensively for making decisions only when stimuli were degraded.

In addition to the significant interaction between polysemy and frequency, the interaction between frequency and stimulus quality was also significant in the combined analysis. The stimulus quality effect was much larger for low-frequency words (195 ms) than for high-frequency words (160 ms). Further, unlike in the lexical decision task, these effects were not modulated by polysemy. Besner and McCann (1987) have also reported a significant interaction between frequency and stimulus quality in a naming task in which stimulus quality was degraded by alternating the case of the stimuli (e.g., *LoSt*). However, their stimulus quality effects were somewhat smaller (35 ms for the low-frequency words and 16 ms for the high-frequency words). Interestingly, in their lexical decision task, they reported the more typical finding of additivity between these factors.

Thus, the implication seems to be that frequency does interact with stimulus quality in naming tasks, regardless of whether the words are ambiguous. Conversely, for the lexical decision task, the interaction only occurs with unambiguous words.

General Discussion

The main purpose of the present research was to determine whether polysemy affects the speed of lexical access and, if so, how. As in Balota and Chumbley (1984, 1985), two word recognition tasks (lexical decision and naming) were assumed to consist of a common lexical access process as well as postaccess, task-specific processes. Any effects common to both tasks would, in the first instance, be considered to be due to the common lexical access process.

In the present studies, the effects of polysemy were evaluated in conjunction with the effects of word frequency. Although both factors affected performance in both tasks, the nature of the relationship between factors was different. Polysemy effects were observed for both high- and low-frequency words in the lexical decision task (Experiment 1), but the effects were limited to the low-frequency words in the naming task (Experiment 2).

Because naming latencies are sensitive to articulation onset differences, a delayed naming task was used to examine the possibility that the different results in the lexical decision and naming tasks were due to articulation onset differences (Experiment 3). The results, however, failed to support this possibility because there was no difference in delayed naming latencies between ambiguous and unambiguous words.

Balota and Chumbley (1984) claimed that naming latencies are better measures of lexical access than lexical decision latencies because the naming task does not require decision-making processes. Thus, naming latencies should be more sensitive to effects occurring during lexical access. On the basis of Balota and Chumbley's claim, one possible explanation for the different effects of polysemy in the two tasks was that polysemy affected lexical access as observed in the naming task but the pattern changed during the decision-making process in the lexical decision task. Another possibility, however, was that the different patterns were each due to the task-specific components. That is, the interaction between polysemy and frequency was due to processes specific to the naming task, and the additive relationship between these variables was due to processes specific to the lexical decision task.

These alternatives were examined in the go–no go naming task (Experiment 4). The former explanation suggests that the results in this task should be the same as those in the lexical decision task because these two tasks are identical except for the modality of response. The latter explanation, however, suggests that there should be an interaction between polysemy and frequency in go–no go naming latencies. The go–no go naming task is assumed to consist of lexical access, decision-making, and pronunciation-related processes. The decision-making process should produce polysemy effects that are additive with frequency. The pronunciation-related process should, however, produce polysemy effects that are interactive with frequency. Therefore, when both processes are combined sequentially in the same task, polysemy effects should appear for both high- and low-frequency words, with the size of the effect being larger for low-frequency words.

The results from the go–no go naming task confirmed the latter predictions. Further, the sizes of polysemy effects and frequency effects were both approximately equal to (but slightly larger than) the sums of the effects in the lexical decision and naming tasks in Experiments 1 and 2. If some portions of the effects in lexical decision and naming were due to lexical access, the sum of the effects in those two tasks should be larger than the effects on go–no go naming latencies because summing those effects should involve adding the lexical access contribution twice. (It is being assumed, of course, that lexical access does not occur twice in the go–no go naming task.) Therefore, these results seem to suggest that both polysemy and frequency have little effect during lexical access but rather these variables affect processes specific to each task.

Most of the previously proposed accounts of polysemy effects have suggested that polysemy affects the lexical access process. Rubenstein et al. (1971) and Jastrzembski (1981) explained polysemy effects by assuming different numbers of lexical entries for ambiguous and unambiguous words. Because ambiguous words have more entries, the speed of accessing one of those entries should be faster. Thus, lexical access would be faster for ambiguous words than for unambiguous words. Balota et al. (1991) also suggested a lexical access account in the framework of the interactive-activation model. They explained the polysemy effect as due to feedback activation from meaning-level units to word-level units. Because ambiguous words have multiple meanings, the feedback activation from meaning-level to word-level units is greater, thus resulting in faster lexical access for ambiguous words. Contrary to these lexical access accounts, however, the present results suggest that there is little influence of polysemy during lexical access.

Some recent results (Kellas et al., 1988) may appear to argue against this conclusion. Participants in these experiments carried out a lexical decision task while at the same time engaging in an auditory probe detection task as a secondary task. The auditory probes were presented following the onset of the lexical decision stimulus by 90, 180, or 270 ms. Polysemy effects were observed not only on lexical decision latencies but also on probe reaction times. That is, the probe reaction times were faster when the lexical decision stimuli were ambiguous words than when the stimuli were unambiguous words. On the basis of these results, particularly the fact that the effects occurred at the very briefest stimulus onset asynchrony (SOA) (90 ms), Kellas et al. suggested that the polysemy effects in both tasks were due to the increased ease of lexical access with ambiguous words.

It is, however, not at all unreasonable to argue that the polysemy effects on probe reaction times were actually due to different attention demands during the decision-making process even at the 90-ms SOA. In particular, it should be noted that the mean probe reaction times were more than 600 ms in the 90-ms SOA condition. Thus, these probe detection responses occurred more than 700 ms after the onset of the lexical decision stimulus. Because the mean lexical decision latencies were 744 ms for ambiguous words and 782 ms for unambiguous words, it seems unlikely that probe processing could have been finished prior to the beginning of the participants' lexical decision-making process, even at the 90-ms SOA. Thus, the probe reaction times were probably also influenced by the difficulty of the decision-making process. As such, these results would also be consistent with a decision-based explanation of polysemy effects, a possibility acknowledged by Kellas et al. (1988): "It may be the decision process itself that is sensitive to the number of meanings, so that ambiguous words speed the postlexical access decision operations and thereby allow attention to be switched sooner" (p. 607).

Word frequency effects have also typically been explained in terms of lexical access operations. The assumption is either that there are frequency-sensitive lexical representations, as in logogen-type models (e.g., Morton, 1969; McClelland & Rumelhart, 1981), or that there is a frequency-ordered matching process to access lexical representations, as in the verification model (Becker, 1980) or the lexical search model (Forster, 1976). Some researchers, however, have recently argued for the possibility of alternative loci for frequency effects. Balota and Chumbley (1984) argued that frequency effects on lexical decision latencies were mostly due to the decision-making processes. Balota and Chumbley (1985) also argued that a part of frequency effects on naming latencies was due to the production process. Further, McCann and Besner (1987) suggested that frequency effects on naming latencies may arise because the connections between orthographic and phonological representations are frequency sensitive.

The present results concerning the effects of frequency seem to be consistent with these more recent accounts. Further, because the identical interaction between polysemy and frequency was observed not only in naming but also in lexical decision with degraded stimuli (Experiment 6), the observed interaction seems not to be attributable to the production process, but rather to arise during the process of retrieving or constructing phonology. It is this process, then, that we are suggesting is the process that is responsible for the effects of both polysemy and frequency (and their interaction) in a naming task.

Operations or Representations?

Given the conclusion that both polysemy and frequency affect task-specific processes, it becomes important to consider the nature of these processes and how they would explain the different types of relationships between polysemy and frequency. Seidenberg and his colleagues (Seidenberg, 1985, 1989; Seidenberg & McClelland, 1989; Waters & Seidenberg, 1985) have argued that one key difference between lexical decision and naming is the types of representations used to accomplish each task.

As demonstrated by Waters and Seidenberg (1985) and in Experiments 5A and 5B, the effects of spelling–sound regularity are, typically, only observed in a naming task (although these effects appear to arise in lexical decision tasks if strange words are included in the stimulus set). Thus, Seidenberg and his colleagues have argued that lexical decisions are usually made based primarily on orthographic information, whereas naming always necessitates the retrieval of phonological information because overt pronunciations are required. They have further argued that although participants often make lexical decisions based primarily on orthographic information, when the orthographic information does not provide enough of a clue to discriminate words from nonwords, participants change their decision strategy to make greater use of phonological information.

Because visual information becomes less available when stimuli are severely degraded, we hypothesized that stimulus degradation would reduce the orthographic familiarity difference between words and nonwords. Thus, with degraded stimuli, word–nonword discriminations should be quite difficult on the basis of orthographic familiarity alone. The important implication was that stimulus degradation should bias participants to make greater use of phonological information in a lexical decision task. The results of Experiment 5C, in particular the existence of a regularity effect for low-frequency words, support the argument that phonological information does play a larger role when stimuli are degraded. On the basis of this result, it was argued that if the interaction between polysemy and frequency were due to the use of phonological representations, the interaction should appear not only on naming latencies but also on lexical decision latencies when the stimuli were degraded. This hypothesis was confirmed. When the stimuli were degraded, the interaction between polysemy and frequency appeared not only in naming (Experiment 7) but also in lexical decision (Experiment 6).

The observed relationships between polysemy and frequency are summarized in Table 12. As this table makes clear, two different patterns of results appeared in the same task (lexical decision) and the identical result (a Polysemy × Frequency interaction) was obtained in both tasks, indicating that the different patterns of relationship between polysemy and frequency were independent of task type. Therefore, it seems quite unlikely that the present results could be explained in terms of the type of operations that were carried out during the task-specific processes. Rather, as argued, these results appear to be due to the nature of the representations used during those processes.

An issue that readers may have noted is that in the present experiments, the observed polysemy effects and the interactions between polysemy and frequency were significant only in the subjects' analyses. Clark (1973) argued that items' analyses, analyses based on treating items as a random factor, are an important component of word recognition research because, unless effects are significant in these types of analyses, investigators are not able to generalize

Table 12
Observed Relationships Between Polysemy and Frequency in Lexical Decision (LDT) and Naming (NM) Tasks

Experiment no.	Task	Stimulus quality	Representations	Experimental results
1	LDT	Clear	Orthographic	Polysemy + Frequency
2	NM	Clear	Phonological	Polysemy × Frequency
6	LDT	Degraded	Phonological	Polysemy × Frequency
7	NM	Degraded	Phonological	Polysemy × Frequency

Note. Four experiments were described in terms of task type, stimulus quality, and the type of representations that seem to play the central role in producing responses. The pattern of the relationships between polysemy and frequency corresponded to the type of representation used to accomplish the task.

their results beyond the particular stimulus set used in that experiment. Thus, the lack of significant effects in our items' analyses may be seen by some as problematic for the present arguments.

On the other hand, however, a number of researchers (Cohen, 1976; Keppel, 1976; Smith, 1976; Wike & Church, 1976) have claimed that Clark's (1973) arguments have limited applicability. To begin with, in word recognition research, the selection of items virtually never is random because researchers are usually attempting to control irrelevant variables during the selection process. Such was clearly the case in the present experiments. The words used here were selected specifically because the word groups could be equated on the extensive set of criteria presented in Table 1. Thus, items was not a random factor in any sense of the word, which, as Wike and Church argue, makes an items' analysis inappropriate. A second point made by Wike and Church is that by treating items as a random factor, the statistical tests have markedly reduced power and, as a consequence, there is much more risk of Type II errors. That is, as in the present experiments, two issues often arise to severely diminish power: (a) the number of stimuli meeting the selection criteria (i.e., the *N* values) is somewhat small, and (b) the important experimental factor or factors must be analyzed in a between- rather than within-item analysis. Wike and Church, in fact, concluded that Clark "is overconcerned with the costs of nonreplicability and underconcerned with the failure to detect differences when they exist" (p. 253).

For these reasons, we have based our theorizing on results from subjects' analyses and not on results from items' analyses. Nonetheless, as noted, for the interested reader, results from items' analyses are reported in the *Results* section for each experiment. In addition, item means for all words are reported in the Appendices for those readers who wish to consider item differences.

Implications of Using Phonological Versus Orthographic Representations

As Fodor (1983) has argued, when reading, visual signals must initially be mapped onto mental representations in order to allow readers to operate further on those representations. According to most word recognition models, these representations are assumed to be word-level representations contained within a lexicon. As a working hypothesis, we have adopted the same assumptions. Within the framework of these models, the notion of *lexical access* has played a central role. As described previously, for example, the word frequency effect has been explained as occurring during the lexical access process by assuming either frequency-sensitive lexical entries or a frequency-ordered serial search mechanism. Also as described previously, polysemy effects have been explained in terms of the lexical access process by assuming multiple lexical representations for ambiguous words. Thus, for these models, the notion of *lexical access* has been crucial because it allowed a number of word recognition phenomena to be accounted for.

The present data, however, suggest that both polysemy and frequency effects are better explained in terms of processes that operate on orthographic and phonological codes. That is, the lexical access process actually appears to play little, if any, role in producing these phenomena. As such, the existence of these effects per se provides little basis for arguing for the existence of a lexical access process. Add to this point the recent arguments by Seidenberg and McClelland (1989) and Seidenberg (1989) challenging the notion of the lexicon itself, and one has to question whether the input process assumed earlier can actually be anything like a *lexical access process.*

Based on a number of these types of considerations, a model like the parallel distributed processing models proposed by Seidenberg and McClelland (1989) and Plaut and McClelland (1993) would seem to provide a better framework for thinking about the present results. In their more general framework these models have separate orthographic, phonological, and semantic levels, each involving distributed representations, although the models implemented when this article was written only contained orthographic and phonological levels.[6] Units at these levels are

---

[6] Because the semantic level was not implemented in Seidenberg and McClelland's (1989) model, any effects that the model predicts would be purely a function of orthographic and phonological factors. Thus, the model actually provides an additional means of evaluating whether our polysemy effects might have been due to orthographic or phonological factors (e.g., spelling–sound regularity). To do this, phonological error scores from Seidenberg and

connected to one another through hidden units. Thus, these models suggest that there would be two independent ways to access word meanings, through orthographic processing or through phonological processing. More important, as in Balota et al.'s (1991) account of polysemy effects, it could be assumed that units at the semantic level feed activation back to units at the other levels. If it is also assumed that ambiguous words have multiple representations at the semantic level (e.g., Balota et al., 1991; Fera et al., 1992), polysemy effects could be accounted for in terms of this feedback process.

More concretely, when processing is based predominantly on orthographic representations (e.g., a lexical decision task with clear stimuli), the process of establishing the orthographic representation to be used in the decision-making process causes the automatic activation of meaning, which then feeds back to influence this orthographically based processing. Similarly, when a task requires phonologically based processing (e.g., a naming task), the process of establishing a phonological representation allows the automatic activation of meaning, which in turn influences that phonologically based processing.

Speaking more generally, the specific architecture that such a model would need to account for the present results would be orthographic input units, orthographic output units, phonological output units, semantic units, and (of course) the hidden units that link the other units. The output units would presumably be the units required for the task-specific processes in the present experiments. That is, in a lexical decision task, although an input stimulus is initially mapped to orthographic input units, participants may tend to base their responses on either the orthographic or the phonological output units. In general, lexical decisions would be made primarily based on the orthographic output activation. When the orthographic output activation does not provide enough of a clue to discriminate words from nonwords, however, phonological output activation must be used. Further, according to the feedback assumption made earlier, the activation at either the orthographic or the phonological levels would automatically spread to semantic units and the semantic units would then send activation back to both the orthographic and the phonological output units. This cascading activation would be greater for ambiguous words because of the summation of activation from multiple semantic units. Thus, the accumulation of output activation would be facilitated for ambiguous words in comparison to unambiguous words and the lexical decision latencies would be faster for ambiguous words.[7]

Similarly, in a naming task, the network computes the phonological output activation from the orthographic input activation. At the same time, semantic units would be activated either from the orthographic input units or from the phonological output units and the activation of semantic units would help to activate the phonological output units by feedback activation. Thus, the accumulation of phonological output activation would be faster for ambiguous words and would lead to faster naming responses.

In a go–no go naming task, lexical decisions would be made primarily based on orthographic output activation whenever the orthographic output activation provides enough of a clue to discriminate words from nonwords. The accumulation of output activation would be facilitated for ambiguous words due to feedback activation from the semantic level, as in a lexical decision task. In addition, after the participant has determined that the stimulus is a word, the phonological output activation would be computed based on the orthographic activation. In theory, this activation could be either orthographic input or orthographic output activation. As in a naming task, this process would be facilitated for low-frequency ambiguous words due to feedback activation from semantic units. The end results would be both an overall polysemy effect and a Polysemy × Frequency interaction, as was observed in Experiment 4.

Assuming the two essentially independent semantic feedback processes from semantic representations, the crucial question is still why we observed the particular relationships between polysemy and frequency reported here. That is, why was the pattern interactive when phonologically based processing was required and additive when orthographically based processing was used?

According to additive factors logic (Sternberg, 1969), an additive relationship between two variables is assumed to indicate that these variables affect separate processing stages, whereas an interactive relationship is assumed to indicate that these variables affect a common stage. Thus, one may wish to argue that polysemy and frequency affect

---

McClelland's model were obtained for our items from a list of error scores for the corpus of words on which the model was originally trained. Because 10 of our items were missing from this list, cell means were used for these items. Mean phonological error scores were 4.03 for low-frequency ambiguous words, 4.46 for low-frequency unambiguous words, 2.94 for high-frequency ambiguous words, and 3.21 for high-frequency unambiguous words. These error scores were then submitted to a 2 (polysemy) × 2 (frequency) ANOVA. The only significant effect was the main effect of frequency, $F(1, 56) = 16.43$, $p < .001$, $MSE = 1.25$. Neither the main effect of polysemy, $F(1, 56) = 1.49$, nor the interaction between polysemy and frequency, $F(1, 56) = .07$, was significant. Because no effect of polysemy was observed in the analysis of phonological error scores, these results provide further evidence that our polysemy effects are truly semantic effects.

[7] Besner, Twilley, McCann, and Seergobin (1990) have shown that Seidenberg and McClelland's (1989) implemented model cannot successfully simulate Waters and Seidenberg's (1985) lexical decision data, raising questions about the model's ability to account for lexical decision data in general. As Besner et al. noted in their Footnote 3, however, it is important to distinguish the implemented model, which does not contain a semantic level, from the model's general framework, in which there is a semantic level. At present, it is far from clear whether the model would still face the same difficulties in explaining lexical decision data if a semantic level were implemented. Thus, although Besner et al.'s point is well taken, it does not appear to cause problems for theorizing based on Seidenberg and McClelland's general framework. In particular, we see no reason that it would cause problems for theorizing about the effects of semantics on lexical decision making.

separate stages in the lexical decision task but affect a common stage in the naming task.

As Sternberg (1969) was careful to point out, however, two factors could affect a common stage in an additive fashion. Thus, the existence of additive effects should only be regarded as evidence for separate stages if the separate stage assumption makes sense in the larger context. Further, because we are attempting to explain our results within a fully interactive network, the additive factors logic must be used a bit cautiously. Thus, in the present instance, the suggestion is that it is more parsimonious to argue that polysemy and frequency affect the same stage but in an additive fashion in the lexical decision task.

Seidenberg and McClelland's (1989) model actually provides an example of how two factors can affect the same stage in an additive fashion. In the model, the effects of regularity and frequency were presumed to occur based on the weights on connections between orthographic input units and phonological output units. Both of these effects would arise during the same processing stage and, thus, this model could simulate the standard interaction between spelling–sound regularity and frequency in a naming task (e.g., Andrews, 1982; Seidenberg et al., 1984; Waters & Seidenberg, 1985). In the model, however, the weights on connections among orthographic input units, hidden units, and phonological output units are adjusted during the training phase, and these weights ultimately depend on the amount of experience with the words themselves and with words that share similar spelling–sound correspondences. What is most relevant is that early in training in their simulation of the regularity effect, the effect appeared not only for low-frequency words but also for high-frequency words. That is, an additive relationship between regularity and frequency was observed. Only with additional training was the regularity effect for high-frequency words reduced and the interactive relationship between regularity and frequency observed.

What occurred in the model is that because high-frequency words were experienced so frequently, the correspondences between spelling and sound for these words became overlearned, and strong connections between a particular spelling and a particular sound were established regardless of the regularity. Thus, the output for both regular and exception words approached asymptote, and regularity had no effect on the computations from orthographic input units to phonological output units for high-frequency words. For low-frequency words, however, the correspondences from a particular spelling to a particular sound were less well learned because of less frequent experience with these words. Thus, for low-frequency words, the connections between a particular spelling and a particular sound were relatively weak. Because of such weak connections, the computations from orthographic inputs to phonological outputs were affected by the spelling–sound correspondences of similarly spelled regular words. Thus, the effects of regularity appeared for low-frequency words.

The important point to be made here is that, applying the same computation from orthographic input to phonological output, Seidenberg and McClelland's (1989) model could produce both additive and interactive relationships between regularity and frequency. This simulation clearly suggests that it is possible for two variables to produce an additive effect even when they affect a common stage. As such, the additive relationship between polysemy and frequency that was observed in the lexical decision task does not necessarily imply that polysemy and frequency affect separate stages. Rather, as suggested by Seidenberg and McClelland's model, the different relationships between polysemy and frequency may reflect a difference in the strength of correspondences between representations, a difference that will change as a function of experience.

Applied to the present results, the argument could be stated as follows: Because the additive relationship between polysemy and frequency appeared when orthographic representations were used, the additive relationship may imply that the connections between orthographic input units and orthographic output units are relatively weak because these links are not used frequently. It is unclear, in fact, what orthographic output units would be used for by a skilled reader, besides making lexical decisions and possibly proofreading. The result would be that feedback from the semantic level could affect processing of both high- and low-frequency words. On the other hand, the interactive relationship between polysemy and frequency was observed when phonological representations were used. Thus, the interaction may suggest that the connections between orthographic input units and phonological output units may be somewhat stronger, especially for high-frequency words. The result would be that a phonological code for high-frequency words could be generated fast enough that feedback from the semantic system would not be able to affect processing, although it would affect processing of low-frequency words. Note that in fact this argument is actually quite similar to the argument that Seidenberg and McClelland (1989) used to explain the development of the interaction between frequency and regularity in their model.

Conclusions

Although Balota et al. (1991) recently suggested the possibility that semantic variables influence the speed of lexical access, the present examination of the effects of polysemy suggested that there is little influence of polysemy during lexical access. Instead, polysemy effects seem to occur during task-specific processes. The additive and the interactive relationships between polysemy and frequency appeared to be due to the type of representations used during the task-specific processes. Note that the claim is not being made that, in all cases, semantic variables do not affect lexical access (assuming, of course, that such a process does in fact exist). Different semantic variables represent different aspects of word meaning and may therefore play different roles. Concreteness, for example, is a variable relating to the contents of word meanings. On the other hand, polysemy is a variable relating to the number of different types of semantic contents. In fact, there are some studies that do seem to suggest that some semantic variables can influence lexical access (e.g., Forster, 1985; Whittlesea & Cantwell, 1987; see also Balota et al., 1991, for a review). Whether these effects will also turn out to have alternative explanations in terms of other processes is a question for future research. Thus, at present, we restrict our conclusions to the polysemy factor. A full understanding of how semantic variables in general are implemented in our language processing system will continue to be an issue of great interest to those trying to understand the reading process.

References

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10,* 565–575.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802–814.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 340–357.

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 and Language, 24,* 89–106.

Balota, D. A., Ferraro, R. F., & Connor, L. T. (1991). On the early influence of meaning in word recognition: A review of the literature. In P. J. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 187–221). Hillsdale, NJ: Erlbaum.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 386–393.

Becker, C. A. (1980). Semantic context effects in visual word recognition: An analysis of semantic strategies. *Memory & Cognition, 8,* 493–512.

Becker, C. A., & Killion, T. H. (1977). Interaction of visual and cognitive effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 389–401.

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* (pp. 201–219). Hillsdale, NJ: Erlbaum.

Besner, D., Twilley, L., McCann, R. S., & Seergobin, K. (1990). On the association between connectionism and data: Are a few words necessary? *Psychological Review, 97,* 432–446.

Brown, P., Lupker, S. J., & Colombo, L. (1994). Interacting sources of information in word naming: A study of individual differences. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 537–554.

Carr, T. H., & Pollatsek, A. (1985). Recognizing printed words: A look at current models. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 1–82). Orlando, FL: Academic Press.

Carr, T. H., Pollatsek, A., & Posner, M. I. (1981). What does the visual system know about words? *Perception & Psychophysics, 29,* 183–190.

Carr, T. H., Posner, M. I., Pollatsek, A., & Snyder, C. R. R. (1979). Orthography and familiarity effects in word processing. *Journal of Experimental Psychology: General, 108,* 389–414.

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

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

Cohen, J. (1976). Random means random. *Journal of Verbal Learning and Verbal Behavior, 15,* 261–262.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic Press.

Coltheart, M., Besner, D., Jonasson, J. T., & Davelaar, E. (1979). Phonological encoding in the lexical decision task. *Quarterly Journal of Experimental Psychology, 31,* 489–507.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). New York: Academic Press.

Cramer, P. A. (1970). Study of homographs. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 367–382). New York: Academic Press.

Fera, P., Joordens, S., Balota, D. A., Ferraro, F. R., & Besner, D. (1992, November). *Ambiguity in meaning and phonology: Effects on naming.* Paper presented at the 33rd Annual Meeting of the Psychonomic Society, St. Louis, MO.

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology.* Cambridge, MA: MIT Press.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales and E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I. (1985). Lexical acquisition and the modular lexicon. *Language and Cognitive Processes, 1,* 87–108.

Forster, K. I., & Bednall, E. S. (1976). Terminating and exhaustive search in lexical access. *Memory & Cognition, 4,* 53–61.

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.

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 674–691.

Jastrzembski, J. E. (1981). Multiple meanings, number of related meanings, frequency of occurrence, and the lexicon. *Cognitive Psychology, 13,* 278–305.

Jastrzembski, J. E., & Stanners, R. F. (1975). Multiple word meanings and lexical search speed. *Journal of Verbal Learning and Verbal Behavior, 14,* 534–537.

Kellas, G., Ferraro, F. R., & Simpson, G. B. (1988). Lexical ambiguity and the time course of attentional allocation in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 601–609.

Keppel, G. (1976). Words as random variables. *Journal of Verbal Learning and Verbal Behavior, 15,* 263–265.

Kliewer, B. D. (1988). *EGA/VGA: A programmer's reference guide.* New York: Intertext/McGraw-Hill.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Mayzner, M. S., & Tresselt, M. E. (1965). Tables of single-letter and digram frequency counts for various word length and letter position combinations. *Psychonomic Science Monograph Supplements, 1,* 13–32.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Ex-*

perimental Psychology: Human Perception and Performance, 13, 14–24.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88,* 375–407.

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1975). Loci of contextual effects on visual word recognition. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 98–118). London: Academic Press.

Millis, M. L., & Button, S. B. (1989). The effect of polysemy on lexical decision time: Now you see it, now you don't. *Memory & Cognition, 17,* 141–147.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165–178.

Nelson, D. L., McEvoy, C. L., Walling, J. R., & Wheeler, J. W. (1980). The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation, 12,* 16–37.

Norris, D. (1984). The effects of frequency, repetition and stimulus quality in visual word recognition. *Quarterly Journal of Experimental Psychology, 36A,* 507–518.

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* (pp. 221–243). Hillsdale, NJ: Erlbaum.

Plaut, D. C., & McClelland, J. L. (1993). Generalization with componential attractors: Word and nonword reading in an attractor network. In *Proceedings of the 15th Annual Conference of the Cognitive Science Society* (pp. 824–829). Hillsdale, NJ: Erlbaum.

Posner, M. I. (1978). *Chronometric explorations of mind.* Hillsdale, NJ: Erlbaum.

Rubenstein, H., Garfield, L., & Millikan, J. A. (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior, 9,* 487–494.

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Homographic entries in the internal lexicon: Effects of systematicity and relative frequency of meanings. *Journal of Verbal Learning and Verbal Behavior, 10,* 57–62.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89,* 60–94.

Seidenberg, M. S. (1985). The time course of information activation and utilization in visual word recognition. In D. Besner, T. G. Waller, & G. E. MacKinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 5, pp. 199–252). Orlando, FL: Academic Press.

Seidenberg, M. S. (1989). Visual word recognition and performance: A computational model and its implications. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 25–74). Cambridge, MA: MIT Press.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96,* 521–568.

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 and Verbal Behavior, 23,* 383–404.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition, 12,* 315–328.

Smith, J. E. K. (1976). The assuming-will-make-it-so fallacy. *Journal of Verbal Learning and Verbal Behavior, 15,* 262–263.

Stanners, R. F., Jastrzembski, J. E., & Westbrook, A. (1975). Frequency and visual quality in a word–nonword classification task. *Journal of Verbal Learning and Verbal Behavior, 14,* 259–264.

Stanovich, K. E., & Bauer, D. W. (1978). Experiment on the spelling-to-sound regularity effect in word recognition. *Memory & Cognition, 6,* 410–415.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling-sound effects in reading: Time course and decision criteria. *Memory & Cognition, 13,* 557–572.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 358–399.

Whittlesea, B. W. A., & Cantwell, A. L. (1987). Enduring influence of the purpose of experiences: Encoding–retrieval interactions in word and pseudoword perception. *Memory & Cognition, 15,* 465–472.

Wike, E. L., & Church, J. D. (1976). Comments on Clark's "The Language-as-Fixed-Effect Fallacy." *Journal of Verbal Learning and Verbal Behavior, 15,* 249–255.

Wilding, J. M. (1988). The interaction of word frequency and stimulus quality in the lexical decision task: Now you see it, now you don't. *Quarterly Journal of Experimental Psychology, 40A,* 757–770.

Appendix A

Ambiguous–Unambiguous Word Pairs Used in Experiments 1, 2, 3, 4, 6, and 7 and Mean Item Response Latencies (in Milliseconds) in Each Experiment

	Experiment 1	Experiment 2	Experiment 3			Experiment 4	Experiment 6	Experiment 7
Item			700	1,000	1,300			
Low frequency–Ambiguous words								
Perch	608	485	421	262	283	769	849	675
Rash	652	479	316	297	267	639	805	748
Punch	540	468	365	258	280	559	662	675
Hail	696	489	362	282	291	721	911	762
Spade	638	480	390	359	356	731	819	667
Shed	650	450	338	320	330	670	869	698
Limp	664	439	334	255	261	685	748	563
Drag	564	474	312	330	275	662	807	663
Seal	590	444	327	318	374	626	712	615
Lean	565	450	348	265	274	606	775	624
Pupil	647	541	349	336	288	711	768	698
Beam	584	460	272	284	313	619	749	685
Bowl	600	486	355	242	247	671	741	673
Sink	590	431	410	314	332	643	725	591
Draft	598	479	289	278	305	593	813	593
M	612	470	346	293	298	660	784	662
Low frequency–Unambiguous words								
Evade	698	507	436	266	272	783	874	803
Cove	607	507	347	395	295	731	765	637
Badge	634	522	301	275	302	687	808	770
Veto	775	555	388	304	297	911	884	758
Sewer	728	477	390	404	317	840	971	696
Wool	685	553	298	278	295	732	860	650
Deaf	596	487	358	270	273	706	794	632
Lung	668	443	281	302	239	632	784	636
Lamp	542	447	311	291	297	573	714	612
Tent	588	490	389	281	282	603	752	659
Solve	562	452	342	358	280	628	712	595
Mode	625	462	289	262	276	672	794	685
Gang	626	479	329	291	271	707	996	673
Pond	583	476	313	311	263	628	723	665
Beard	593	494	309	305	302	715	843	809
M	634	490	339	306	284	703	818	685
High frequency–Ambiguous words								
Well	600	525	288	251	297	627	742	663
Shot	522	415	377	302	335	548	730	569
Watch	581	496	279	259	275	627	734	582
Fine	535	465	296	309	292	551	698	602
March	580	468	294	296	258	618	788	641
Miss	544	438	298	274	322	567	768	594
Mass	570	441	280	251	289	586	713	603
Pass	511	451	325	272	326	543	645	600
Base	557	465	342	269	261	577	671	627
Date	538	442	305	300	313	574	699	671
Post	570	457	332	332	257	590	725	610
Order	534	451	316	290	237	554	694	666
Club	522	461	325	293	339	589	654	580
Range	560	472	387	245	242	591	798	657
Right	479	417	312	294	293	549	663	614
M	547	458	317	282	289	579	715	619

(*Appendix continues on next page*)

Appendix A (continued)

Item	Experiment 1	Experiment 2	Experiment 3			Experiment 4	Experiment 6	Experiment 7
			700	1,000	1,300			
High frequency–Unambiguous words								
Also	561	441	333	257	292	577	702	566
Clay	584	491	323	315	281	639	720	608
Event	567	463	312	282	253	613	699	661
Food	538	470	352	280	309	581	659	600
Green	556	483	307	315	307	553	708	599
Half	559	446	318	279	342	613	733	659
Lack	605	438	321	275	271	675	736	552
Lady	546	449	357	248	258	556	658	586
Loss	520	432	355	311	243	571	617	569
News	595	479	286	254	277	628	814	642
Nine	582	477	287	253	268	622	801	647
Often	562	442	342	272	263	602	752	607
Paid	545	464	309	264	346	588	730	711
River	546	425	294	305	308	587	701	642
Small	546	475	348	358	336	530	673	542
M	561	458	323	285	290	596	714	613

Appendix B

Regular and Exception Words Used in Experiments 5A, 5B, and 5C and Mean Item Response Latencies (in Milliseconds) in Each Experiment

Item	Experiment 5A	5B	5C	Item	Experiment 5A	5B	5C
Low frequency–Regular words				High frequency–Regular words			
Wade	487	633	782	Nine	432	541	698
Peel	455	514	709	Wall	477	549	706
Wick	460	685	757	Book	431	474	639
Sock	488	557	701	Soon	490	507	668
Wink	460	616	718	Feel	475	505	629
Dock	459	549	722	Help	423	465	626
Dusk	450	562	752	Best	436	480	617
Rust	439	523	688	Face	447	477	605
Sank	509	623	776	Less	432	521	662
Beam	485	531	758	Take	462	486	643
M	469	579	736	*M*	451	501	649
Low frequency–Exception words				High frequency–Exception words			
Wand	474	645	850	Foot	458	489	631
Crow	536	599	794	None	425	510	654
Worm	467	521	711	Move	439	483	660
Pear	555	580	705	Love	427	474	591
Lure	497	684	819	Sure	523	485	654
Pour	508	576	778	Word	450	502	642
Doll	475	576	757	Done	458	499	639
Warn	491	544	819	Want	432	504	713
Pint	564	599	742	Give	447	487	640
Root	470	509	666	Good	452	485	638
M	504	583	764	*M*	451	492	646

Received June 23, 1993
Revision received June 22, 1995
Accepted August 15, 1995 ■