UNIVERSITY OF CALGARY

Does Print Exposure Modulate the Effect of Word Frequency and the Effect of Neighborhood Size?

by

Verna Chow

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

August 1, 2001

0-612-65096-0

Abstract

The purpose of this study was to examine whether print exposure modulates the effect of word frequency and the effect of neighborhood size, given the conflicting finds of Lewellen et al. (1993) and Chateau and Jared (2000). In Experiment 1, like the Lewellen et al. experiment, regular nonwords were used in the lexical task. There was no evidence that print exposure had any effect on the responses to words or to nonwords. This result was confirmed in Experiment 3. In Experiment 2, like the Chateau and Jared experiment, pseudohomophones were used in the lexical decision task. In this experiment there was clear evidence of an affect of print exposure on the responses to words and to pseudohomophones. In particular, participants with lower levels of print exposure showed larger word frequency effects and larger neighborhood size effects than participants with higher levels of print exposure, and they were also slower to reject pseudohomophones. Several analyses revealed that the low print exposure participants were slower and more error prone in Experiment 2 than in Experiment 1, and, in addition, exhibited a larger word frequency effect and a larger neighborhood size effect in Experiment 2. Their responses to the pseudohomophones used in Experiment 2 were also slower and more error prone than their responses to the regular nonwords used in Experiment 1. Overall, there was unmistakable evidence that the low print exposure participants responded differently than both the high print exposure participants in Experiment 2 and the low print exposure participants in Experiment 1. In contrast, the responses of the high print exposure participants in both experiments were virtually identical. The implications of these findings for understanding the affect of print exposure on word identification are discussed.

Acknowledgements

I am grateful to have been mentored and to have learned a great deal about research in cognitive psychology from my supervisor Chris Sears, who I would wish to thank for motivating me to higher achievements, for believing in me, for his help, brilliant creativity, and enthusiasm to guide me through this thesis from its conception to its completion. I would also like to thank Penny Pexman and John Mueller for their time and help with this research project, and to Susan Boon and Michelle Jacobsen for serving on my examination committee.

I am also grateful for the friendship and support of Andrea Welder, Allisa Pencer, Phillip Litke, Audny Dypvick, Lenora Brown, Alisa McArthur, Geoff Ho, Pam Valentine, Collin David, Karen Soriano, Luca Ermini and Shirley Mendosa. I thank Karla Gordon for all the interesting conversations and for her brilliant ideas from which I will always benefit. As well, I thank the rest of the Gordon clan, John, Caleb and Josh for their continued support and appreciation. I thank Amy Deller for her encouragement and positive spirit. I thank Crystal Sharp for her feedback and advice. I thank Lori Bednarick for the long and interesting conversations we had until dawn and her inspired support. Finally, I wish to thank my parents for their support during the last two years.

TABLE OF CONTENTS

LIST OF TABLES

Does Print Exposure Modulate the Effect of Word Frequency and the Effect of Neighborhood Size?

It has been known for some time that large individual differences in adult reading achievement exist, even among college students. Perfetti (1985), for example, noted that the reading speeds of college students can be as high as 400 words per minute and as low as 150 words per minute. In the past few decades a number of studies have sought to delineate the individual differences in orthographic and phonological processing skills thought to underlie individual differences in reading ability.

Phonological processing skills have been shown to vary as a function of reader skill in experimental studies. Chateau and Jared (2000), for example, examined phonological processing ability in a pseudoword naming task, and phonological and orthographic processing in a form priming task (Forster, 1987; Forster & Davis, 1991). In the pseudoword naming task, participants were presented with pronounceable pseudowords (e.g., SHUP) and were asked to read them aloud as quickly and as accurately as possible. The higher skilled readers read the pseudowords more quickly and made fewer errors than the lower skilled readers, which suggests that the higher skilled readers activated phonological representations faster than the lower skilled readers (identical results were reported by Stanovich & West, 1989; Bell & Perfetti, 1994, and Jared, Levy, & Rayner, 1999). Bell and Perfetti also reported that lower skilled readers were slower than higher skilled readers to pronounce low-frequency words, but not high-frequency words. On the critical trials in the form priming lexical decision task, participants saw a very briefly presented (30 ms or 60 ms) prime word, and then a target that was orthographically similar but phonologically dissimilar to the prime (e.g., TOUCH-COUCH). Relative to the trials in which the prime and target were completely unrelated (e.g., SHALL-COUCH), for the higher skilled readers, responses to COUCH were facilitated after a 30 ms prime presentation, but were delayed after a 60 ms prime presentation. But this was only true for the higher skilled participants. The lower skilled readers' responses to the target words were not affected by the type of prime word or its duration. Chateau and Jared concluded that higher skilled readers quickly activate an orthographic representation of the prime word (which facilitates responses to an orthographically similar target), and shortly thereafter a phonological representation of the prime is even more strongly activated (which delays responses to a phonologically dissimilar target). According to Chateau and Jared, the absence of this pattern of effects for the lower skilled readers suggests that they are slower to activate either type of representation, or that both are equally activated during the intervals examined (30 ms and 60 ms).

Orthographic processing skills have also been shown to vary as a function of reader skill in experimental studies. Butler, Jared, and Hains (1984), for example, found substantial differences among adult readers in the use of orthographic information. In their Experiment 1 (a tachistoscopic recognition task), participants were presented with a series of zero order (i.e., random letter sequences), second order, and fourth order (e.g., DINGLECT) approximation-to-English letter strings and were asked to report the order in which the letters were presented. For the zero order sequences, the high vocabulary participants (as assessed by the Nelson-Denny Reading Test; Brown, 1900) recalled more of the left-most letters, and for the second and fourth order sequences more letters from the right of the string were recalled. The low vocabulary participants, in contrast, recalled more letters from the right end of the string for all three types of sequences. High vocabulary participants were also found to make greater use of orthographic knowledge by parsing the letter sequences into larger units, thereby recalling more of the letter strings (and also, as a result, being less affected by word length).

In a second experiment, eight-letter words were presented sequentially in three separate units, which either matched or violated syllable boundaries in the word (e.g., HO/SPITA/L; HOS/PI/TAL; HO/SP/ITAL; HOS/PIT/AL). Participants were asked to report the word verbally, to guess if uncertain, or to report any letters they had seen if they were unable to report an entire word. High vocabulary participants were sensitive to the type of unit shown and made more errors when words were divided into non-syllabic units (e.g., HO/SPITA/L). Low vocabulary participants, in contrast, made a similar percentage of errors whether the presentation matched or violated syllable boundaries. Butler et al. suggested that this indicated that the low vocabulary participants were not as sensitive to the type of unit presented (syllable or non-syllable).

In a study by Bell and Perfetti (1994), skilled and less skilled readers were compared on a number of information processing and language comprehension tasks that estimate cognitive components of reading. Reader skill was assessed using the Nelson-Denny Reading test (1981), SAT scores (College Board Scholastic Aptitude Test, 1971), and a spatial ability test (an approximation of IQ performance). Bell and Perfetti administered two variants of a lexical decision task. In the orthographic choice task, participants had to decide which of two letter strings was a real English word (e.g., BEAR or BAIR). In the phonemic choice task, participants had to decide which of two pseudowords could be pronounced as a real word (e.g., BAIR or BOIR). Skilled readers were faster and more accurate on both of these tasks. According to Bell and Perfetti, lower skilled readers have fewer accessible representations of words, not merely less effective orthographic-phonemic processing skills. If phonemic decoding were the only problem in word identification for the lower skilled readers, then it would be expected that only the phonemic choice task would produce differences.

In Bell and Perfetti's (1994) spelling task, an experimenter read words aloud and participants were asked to spell these words. The spelling test consisted of "common" words (high-frequency words), "familiar" words (low-frequency words), and "unfamiliar" words (low-frequency words judged by the experimenters to be infrequently encountered). Half of these words were regular words (e.g., CONGRESS) and the remainder were irregular (e.g., ISLAND). In addition, there was a group of unfamiliar words with regular spelling patterns that, in the experimenters' opinion, most participants had probably never seen or heard (e.g., LACHRYMOSE). Not surprisingly, overall, the high skilled readers were better spellers than the low skilled readers. But large differences between the two groups were evident only when the words were of low-frequency or were unfamiliar, particularly when the words were irregular. According to Bell and Perfetti, their data suggest that low skilled readers possess lower quality orthographic representations with fewer connections between spelling patterns and phoneme sequences.

Reader Skill and the Role of Print Exposure

One variable that is known to contribute to the differences between high and low skilled readers is the degree of print exposure. That is, higher skilled readers appear to read more often than lower skilled readers, and as such spend more time building their vocabulary and further refining their orthographic and phonological processing abilities. Bräten, Lie, Andreassen, and Olaussen (1999) found that, for children, the amount of leisure time spent reading directly increased orthographic processing skills, independently of phonological processing skills. Several studies (Chateau & Jared, 2000; Lewellen, Goldinger, Pisoni, & Greene, 1993; Stanovich & West, 1989) have also shown that print exposure is related to individual differences in adult reader skill. As a consequence, there is growing interest in the role that print exposure plays in reading processes.

Until fairly recently, the most common method used to assess print exposure was through the use of questionnaires, in which children and adults are asked to estimate the amount of time they spend reading. This technique has been found to be unreliable, however, because of a tendency for participants to represent themselves in a socially desirable way by overestimating their print exposure (Stanovich & West, 1989; West, Stanovich, & Mitchell, 1993). Stanovich and West (1989) developed a measure of print exposure that is both statistically reliable and less subject to social desirability effects. The Author Recognition Test (ART) is a checklist of 100 names. A participant is asked to place a checkmark next to the names they know to be an author or a writer. The list consists of the names of 50 popular authors and 50 foils. Many of the authors regularly appear on best seller lists, and are from a wide variety of genres (e.g., humor, religion, biography, mystery, romance, etc.). Guesses are taken into account by subtracting incorrect responses from correct responses when calculating an overall score.

Stanovich and West (1989) investigated the relation between print exposure (as measured by the ART) and orthographic processing skills. In Experiment 1, a spelling test was used to assess linguistic sensitivity. The spelling test consisted of three sets of words: an orthographic set, a morphophonemic set, and an exception set. The orthographic set consisted of words that had ambiguous segments, which could be spelled by the application of established orthographic patterns or by regularities at the surface phonetic level. For example, the orthographic set required participants to decide whether or not to drop the e before a suffix beginning with a vowel (e.g., CHANGEABLE). Words in the morphemic set required participants to draw upon abstract morphophonemic knowledge to derive spelling conventions, such as the knowledge that one must double the r in CONFERRING but not in CONFERENCE. The exception set consisted of words with sound-spelling relations that occur very infrequently in English (e.g., SERGEANT). In the experiment, the experimenter pronounced each word, used it in a meaningful sentence, and then pronounced it again. Participants were then asked to print the spelling of the word. Stanovich and West (1989) reported that ART scores were correlated with performance on this task ($r = .46$). Specifically, participants with higher ART scores (and thus higher levels of print exposure) performed better on the spelling test than participants with lower ART scores.

In a second experiment, Stanovich and West (1989) examined whether individual differences in print exposure were related to orthographic processing skills, once phonological processing skills had been taken into account. Orthographic processing skills were measured with an orthographic choice task and a homophone choice task. In the orthographic choice task, participants viewed a word/nonword pair that sounded alike (e.g., ROOM/RUME) and indicated which one was a real word. In the homophone choice task, a question was read to participants (e.g., Which is a fruit?) and to answer they chose between two homophones presented on a computer screen (e.g., PEAR/PAIR). Two types of tasks were used to tap phonological processing, the phonological choice task and the pseudoword naming task. In the phonological choice task, participants viewed pairs of pseudowords (e.g., KAKE/DAKE) and then chose the one that sounded like a real word. In the pseudoword-naming task, the stimuli were pronounceable nonwords (e.g., MURSH) which the participants were asked to read aloud.

Stanovich and West (1989) reported that ART scores correlated with performance on the orthographic processing tasks and on the phonological processing tasks. Greater print exposure, as measured by the ART, was associated with superior orthographic and phonological processing skills. A multiple regression analyses indicated that the ART was a significant predictor of performance on the orthographic processing tasks even after phonological processing ability was partialed out. Stanovich and West concluded that print exposure contributes to the efficiency of orthographic processing independently of its contribution to phonological processing skills.

Modulation of Word Frequency and Neighborhood Effects by Print Exposure

Many studies have indicated that greater print exposure is associated with superior orthographic and phonological processing skills. Nonetheless, the vast majority of word identification research ignores individual differences in print exposure (see Jared et al., 1999, and Pexman, Lupker, & Reggin, 2001, for notable exceptions). Two recent experiments have specifically examined the relation between print exposure and the speeded identification of single words (Lewellen, Goldinger, Pisoni, & Greene, 1993; Chateau & Jared, 2000). These two experiments sought to determine if print exposure modulates the effect of word frequency and the effect of neighborhood size in a lexical decision task.

The word frequency effect is probably the most statistically robust finding in word recognition research. Words that are encountered very frequently are recognized more rapidly than words that are encountered infrequently (Forster & Chambers, 1973). Word frequency is quantified using printed frequency counts such as the Kucera and Francis (1967) normative frequencies. The word frequency effect is quantified as the difference in performance (e.g., response latencies, error rates) between high-frequency words and low-frequency words. The word frequency effect is thought to occur because the orthographic representations of words that are encountered very frequently are more accessible than the representations of words that are encountered infrequently.

The neighborhood size effect refers to the fact that responses to words with large orthographic neighborhoods are faster and are more accurate than responses to words with small orthographic neighborhoods (a facilitatory neighborhood size effect; e.g., Andrews, 1989, 1992; Forster & Shen, 1996; Sears, Hino, & Lupker, 1995). A word's orthographic neighborhood is defined as the number of words that can be created by changing one letter of the word while maintaining letter positions (Coltheart, Davelaar, Jonasson, and Besner, 1977). For example, the orthographic neighbors of the word FACT are FACE, FAST, PACT, and TACT. The neighborhood sizes of words varies considerably, with some words (e.g., MALE) having more than twenty neighbors and others having few neighbors or none at all (e.g., GIRL with two neighbors; IDOL, with no neighbors). Importantly, the facilitatory neighborhood size effect is usually only observed for low-frequency words. Like the word frequency effect, the neighborhood size effect is thought to be an orthographic effect due to orthographic processing. That is, a reader's lexicon stores the orthographic patterns of thousands of words, and during the process of word identification the existence of words similar to the presented word facilitate its identification (Andrews, 1989, 1992; Sears et al., 1995; Sears, Lupker, & Hino, 1999; Sears, Hino, & Lupker, 1999; Siakaluk, Sears, & Lupker, 2001).

Lewellen et al. (1993) predicted that the neighborhood size effect would be larger for participants with high levels of print exposure than for participants with low levels of print exposure. They reasoned that, for the high print exposure participants, because of their larger vocabulary, more of the neighbors of large neighborhood words would be activated, enhancing the facilitation produced by large neighborhoods. They also reasoned that the effect of word frequency would be smaller for high print exposure participants because of their greater exposure to low-frequency words. That is, for the high print exposure participants, the experiential difference between high-frequency words and low-frequency would be smaller than for low print exposure participants, which would produce a smaller word frequency effect in both response latencies and in error rates.

Participants were assigned to high print and low print exposure groups on the basis of word familiarity ratings, vocabulary test scores, and responses to a language experience questionnaire. Those who scored above the median on all three measures were assigned to the high print exposure group and those who scored below the medians were assigned to the low print exposure group. Two additional paper and pencil tasks were also administered, the ART and the MRT (i.e., the Magazine Recognition Test, which measures recognition of magazine titles). As expected, the high print exposure participants identified more authors and magazine titles than the low print exposure participants.

In Lewellen et al.'s (1993) lexical decision experiment, word frequency and neighborhood size were factorially manipulated; the words were of low or of high frequency, and had few neighbors or had many neighbors. There were two blocks of trials in the experiment. In one block of trials the nonwords were orthographically illegal and unpronounceable (e.g., LPREE), and in a second block of trials the nonwords were orthographically legal and pronounceable (e.g., THRAD). The high and low print exposure groups only differed statistically from one another in the latter case. In this block, the high print exposure participants made fewer errors to the nonwords and rejected them more quickly than the low print exposure participants. Responses to words were also slightly faster and less error prone for the high print exposure participants, but neither the effect of word frequency nor the effect of neighborhood size interacted with print exposure. When the identical word stimuli were used in a pronunciation task similar results were obtained. Because the two groups of participants exhibited similar word frequency and neighborhood size effects, Lewellen et al. (1993) concluded that the effect of word frequency and the effect of neighborhood size were not modulated by individual differences in print exposure.

Like Lewellen et al. (1993), Chateau and Jared (2000) expected the word frequency effect would be smaller for participants with high levels of print exposure, because high print exposure participants would encounter low-frequency words more often than low print exposure participants. Unlike Lewellen et al. (1993), Chateau and Jared (2000) predicted that participants with high levels of print exposure would exhibit a smaller neighborhood size effect than participants with low levels of print exposure. That is, according to Chateau and Jared's reasoning, because the neighborhood size effect is smaller for high-frequency words than for low-frequency words (Andrews, 1997), and because high print exposure participants would encounter any given set of words more often than low print exposure participants, they should be less affected by experimental manipulations of neighborhood size. Consistent with both of these predictions, they reported a word frequency effect of 128 ms for their high print exposure participants and 278 ms for their low print exposure participants, and neighborhood size effects of 58 ms and 112 ms, respectively. Thus, in contrast to the results of Lewellen et al. (1993), Chateau and Jared's (2000) results suggest that both the neighborhood size effect and the word frequency effect are modulated by individual differences in print exposure.

The Present Study

The purpose of the present study was to re-examine whether the word frequency effect and the neighborhood size effect are modulated by print exposure. The results of Lewellen et al. (1993) and Chateau and Jared (2000) lead to very different conclusions, and this empirical contradiction is exacerbated by a number of important differences between these two studies.

First, the interpretation of Chateau and Jared's (2000) data is complicated by the fact that error rates were very high for the low-frequency words used in their experiment. Error rates exceeded 30% in some conditions, even for their high print exposure participants. In contrast, the error rates of Lewellen et al.'s high print exposure participants did not exceed 5%. This suggests that many of their words were not known to their participants, and so the response latency data may not reflect normal word identification processes. Second, whereas Chateau and Jared used the ART to assign participants to groups, Lewellen et. al. did not (although it was found to correlate with their selection measures). Moreover, Chateau and Jared equated their high and low print exposure groups on Nelson-Denny reading comprehension scores, whereas Lewellen et al. made no effort to equate their two groups on any measure of language proficiency. As Chateau and Jared noted, when reading comprehension ability is equated, the variation in word recognition processes left to be accounted for by print exposure is greatly reduced. As a consequence, their study was a very conservative test of the effects of print exposure on word recognition. Third, Lewellen et al. used pronounceable nonwords (e.g., BRINT) in their lexical decision task, whereas Chateau and Jared used pseudohomophones, nonwords that sound like words when pronounced (e.g., BRANE). When pseudohomophones are used in a lexical decision task, participants should rely more on orthographic information than on phonological information when making their word/nonword decisions, because phonological information cannot be used to distinguish the words from the nonwords. Indeed, according to some investigators, under these conditions the use of phonology should be greatly reduced or even eliminated (e.g., Davelaar, Coltheart, Besner, Jonasson, 1978; McQuade, 1981, 1983; Monsell, Patterson, Graham, & Hughes, 1992; Pugh, Rexer, & Katz, 1994). As a consequence, the participants in Chateau and Jared's study may have relied on orthographic information more than the participants in Lewellen et al.'s study. Any and all of these differences could have contributed to the very different outcomes observed in these two experiments.

Given the number of differences between the Lewellen et l. (1993) and Chateau and Jared (2000) experiments, it is not feasible to assess the impact of each of them in a single experiment. The strategy taken in this study was to focus first on the different nonword manipulations used. To this end, in Experiment 1, the effect of print exposure on the word frequency effect and the neighborhood size effect was assessed when regular nonwords (e.g., BRINT) were used in the lexical decision task. In Experiment 2, the same word stimuli were presented to a new group of participants (again divided into high and low print exposure groups) and a set of pseudohomophones was used as the nonword stimuli. An effect of print exposure in one experiment and not the other would indicate that the type of nonword is critical for observing a print exposure effect. On the other hand, if both experiments produced the same outcome, then other differences between the Lewellen et al. and the Chateau and Jared experiments must be responsible for their conflicting findings. One obvious possibility is whether the print exposure groups are equated on reading comprehension ability (as they were in Chateau and Jared's experiment). It was, however, not necessary to evaluate this possibility until the effect of nonword type was first evaluated, and so in Experiments 1 and 2, no effort was made to equate the high print and low print exposure groups on measures of reading comprehension ability. Finally, to avoid the problems inherent with interpreting response latencies when error rates are very high, in both experiments efforts were made to keep the error rates fairly low. This was accomplished by ensuring that no unusual or rare words were used in the experiments, by instructing participants to give preference to accuracy over speed, and by providing error feedback (percent error) after every block of 33 trials.

Experiment 1

Like the Lewellen et al. (1993) and Chateau and Jared (2000) experiments, the purpose of this experiment was to examine the relation between print exposure, the word frequency effect, and the neighborhood size effect. To this end, in this experiment word frequency and neighborhood size were factorially manipulated, and participants were divided into a high print exposure group and a low print exposure group using the ART. Like the Lewellen et al. (1993) experiment, the nonwords used in the lexical decision task were orthographically legal and pronounceable, but were not pseudohomophones. Unlike the Lewellen et al. experiment, the neighborhood sizes of the nonwords was also manipulated to assess for any effect of print exposure on nonword response latencies and error rates. Responses to nonwords with large neighborhoods (e.g., SORK) are slower and more error prone than responses to nonwords with small neighborhoods (e.g., AVAR). This inhibitory neighborhood size effect is due to the fact that nonwords with many neighbors are more word-like than nonwords with few neighbors, and so are more difficult to reject in a speeded decision task (e.g., Andrews, 1989, 1992, Forster & Shen, 1996; Sears et. al., 1995). By manipulating the nonword neighborhood size in this experiment, any influence of print exposure on this aspect of orthographic processing could also be ascertained.

Method

Participants. One-hundred and twenty University of Calgary undergraduate students participated in this experiment. All participants received bonus credit toward a psychology course in exchange for their participation. The age of the participants ranged from eighteen to forty-five years of age. All participants were native English speakers and reported normal or corrected to normal vision.

Participants were administered a Canadian version of the ART (Chateau & Jared, 2000) at the start of the experimental session. For the data analyses, two groups of participants were formed using a tertile split of the ART scores. Thirty-eight participants were assigned to the high ART group (with a mean ART score of 26.4) and thirty-seven were assigned to the low ART group (with a mean ART score of 10.0). Note that the mean ART scores of these two groups were very similar to the mean ART scores of Chateau and Jared's high ART and low ART groups (22.3 and 11.3, respectively), and Lewellen et al.'s high and low print exposure groups (29.8 and 13.25, respectively).

Stimuli. The descriptive statistics for the word stimuli are listed in Table 1. Two factors were manipulated. The first factor was printed word frequency. Half of the words were high-frequency words with a Kucera and Francis (1967) mean normative frequency per million words of 189.9 (range of 70 to 683), and the remainder of the words were of low-frequency with a mean normative frequency of 22.6 (range of 3 to 49).

The second factor manipulated was neighborhood size. Half of the words had a small neighborhood (i.e., at least one neighbor and no more than 5 neighbors); these had a mean neighborhood size of 2.7. The other half of the words had a large neighborhood (i.e., at least 7 neighbors; range of 7 to 18); these had a mean neighborhood size of 10.5. To be considered a neighbor of a target word, a word had to appear either in the Kucera and Francis (1967) norms or in an 80,000 word computer-based dictionary. There were 33 words in each of the four conditions. Within each condition there were 18 four-letter words and 15 five-letter words.

All of the nonwords were orthographically-legal and pronounceable, and were matched closely to the words in neighborhood size. More specifically, for the small neighborhood nonwords the mean neighborhood size was 2.8 (range of 1 to 5), and for the large neighborhood nonwords the mean neighborhood size was 10.9 (range of 6 to 17). (The overall mean neighborhood size of the nonwords was 6.8.) The nonwords were also matched to the words on length, such that 72 were four letters in length and 60 were five letters in length. There were 66 nonwords with small neighborhoods and 66 nonwords with large neighborhoods. Thus, 264 stimuli were presented in the experiment, 132 words and 132 nonwords.

Apparatus and Procedure. Stimuli were presented on a color VGA monitor driven by a Pentium-class microcomputer. The presentation of stimuli was synchronized with the vertical retrace rate of the monitor (14 ms) and response latencies were measured to the nearest ms. At a viewing of 50 cm the stimuli subtended a visual angle of approximately 1.1 degrees.

Each trial was initiated by a 1 2000 Hz warning tone, after which a fixation point appeared at the center of the video monitor. The fixation point was presented for 1 s, and then was replaced by a word or nonword stimulus (presented in uppercase letters). Participants indicated the lexicality of stimuli (word or nonword) by pressing one of two buttons on a response box. The participant's response terminated the stimulus display and the next trial was initiated after a timed interval of 2 s. The order in which the experimental stimuli were presented was randomized separately for each participant.

Each participant completed 16 practice trials prior to the collection of data. The practice stimuli consisted of eight words (four of low-frequency, four of high-frequency) and eight orthographically-legal and pronounceable nonwords. Four of the nonwords had a small neighborhood and four had a large neighborhood. (These practice stimuli were not used in the experiment, and the data from these practice trials were not analyzed.) Following the practice trials the participants were provided with feedback as to the mean latency and accuracy of their responses (percent error), and during the experimental trials this information was presented every 33 trials. Participants were instructed to respond as quickly as possible, while keeping their error rate below 5%.

Results

For the word data, response latencies and error rates were submitted to a 2 (Group: high ART, low ART) x 2 (Word Frequency: high, low) x 2 (Neighborhood Size: small, large) mixed-model factorial analysis of variance (ANOVA). For the nonword data, response latencies and error rates were submitted to a 2 (Group: high ART, low ART) x 2 (Neighborhood size: small, large) mixed-model ANOVA. Both subject (F_s) and item (F_i) analyses were carried out. Response latencies to words and to nonwords that were less than 250 ms or greater than 1500 ms were considered outliers and were removed from the data set. A total of 61 observations (0.30% of the data) were removed by this procedure (30 response latencies of the high ART participants and 31 response latencies of the low ART participants).

Response Latencies for Words. The mean response latencies of correct responses and the mean error rates are shown in Table 2. The main effect of group was not significant (both F's < 1), as the average response latency for the low ART group and for the high ART group was virtually identical (538 ms and 539 ms, respectively). The main effect of word frequency was significant, $F_s(1, 73) = 259.0$, $p < .001$, $MSE = 445.8$; $F_i(1, 256) = 74.6$, $p < .001$, $MSE = 1{,}439.1$. Responses to high-frequency words were an average of 39 ms faster than responses to low-frequency words. The main effect of neighborhood size was also significant, $F_s(1, 73) = 24.8$, $p < .001$, $MSE = 368.7$; $F_i(1, 256) = 6.19$, $p < .05$, $MSE = 1{,}439.1$. Responses to words with large neighborhoods were an average of 11 ms faster than responses to words with small neighborhoods. As expected, the neighborhood size effect interacted with word frequency, $F_s(1, 73) = 17.3$, $p < .001$, $MSE = 245.2$; $F_i(1, 256) = 2.96$, $p = .08$, $MSE = 1{,}439.1$. For the low-frequency words, words with large neighborhoods were responded to an average of 18 ms faster than words with small neighborhoods, $F_s(1, 73) = 37.6$, $p < .001$, $MSE = 343.5$; $F_i(1, 128) = 6.33$, $p < .05$, $MSE = 2{,}012.6$. For the high-frequency words there was no neighborhood size effect (both p's $> .10$).

The interaction between group and word frequency was not significant, $F_s(1, 73) = 2.05$, $p > .15$, $MSE = 445.8$; $F_i < 1$, nor was the interaction between group and neighborhood size (both F's < 1). The high ART and the low ART participants exhibited very similar word frequency effects (35 ms and 43 ms, respectively), and almost identical neighborhood size effects (10 ms and 11 ms, respectively). The three-way interaction between group, word frequency, and neighborhood size was also not significant (both F's < 1). As can be seen in Table 2, in each condition the mean response latencies of the two groups were almost indistinguishable from one another.

The absence of an interaction between group and word frequency was not due to a weak manipulation of word frequency, and a consequent reduced statistical power to detect this effect. Although the word frequency manipulation was stronger in Lewellen et al.'s experiment (i.e., the normative frequencies of their low- and high-frequency words were 6.6 and 683.3, respectively; in Experiment 1 they were 22.6 and 189.9, respectively), the word frequency effect was larger in this experiment (39 ms) than it was in Lewellen et al.'s experiment (21 ms). Moreover, a stronger word frequency manipulation in a posthoc analysis produced a larger word frequency effect, but still no Group x Word Frequency interaction (both p's $> .10$). In this analysis, the mean normative frequency of the low-frequency words was 10.2 (N = 28), and the mean normative frequency of the high-frequency words was 303.8 (N = 28). The overall word frequency effect was 65 ms. For the low print exposure group the word frequency effect was 72 ms; for the high print exposure group it was 57 ms. (In the same analysis, for the low-frequency words, the neighborhood size effect was 40 ms for the high ART participants and 33 ms for the low ART participants, and again there was no Group x Neighborhood Size interaction.)

Error Rates for Words. The main effect of group was marginally significant, $F_s(1, 73) = 3.22$, $p = .07$, $MSE = 20.6$; $F_i(1, 256) = 3.19$, $p = .07$, $MSE = 18.3$. Overall, the low ART participants made slightly more errors than the High ART participants (3.9% vs. 3.0%). As was the case in the response latency analysis, there were significant main effects of word frequency, $F_s(1, 73) = 54.5$, $p < .001$, $MSE = 11.9$; $F_i(1, 256) = 31.3$, $p < .001$, $MSE = 18.3$, and of neighborhood size, $F_s(1, 73) = 9.34$, $p < .01$, $MSE = 11.1$; $F_i(1, 256) = 4.99$, $p < .05$, $MSE = 18.3$. Participants made fewer errors to high-frequency words (2.0%) than to low-frequency words (4.9%), and fewer errors to words with large neighborhoods (2.8%) than to words with small neighborhoods (4.0%). The interaction between word frequency and neighborhood size was also significant, $F_s(1, 73) = 12.0$, $p < .01$, $MSE = 8.59$; $F_i(1, 256) = 4.95$, $p < .05$, $MSE = 18.3$. For the low-frequency words, participants made fewer errors to words with large neighborhoods than to words with small neighborhoods (3.7% vs. 6.1%), whereas for the high-frequency words there was no effect of neighborhood size (2.0% for the words with large neighborhoods and 2.0% for the words with small neighborhoods).

The interaction between group and word frequency was not significant (both F's < 1), nor was the interaction between group and neighborhood size (both p's > .20), nor was the three-way interaction (both F's < 1). Thus, apart from the small difference in error rates between the two groups, the error analysis produced the same results as the response latency analysis; namely, no evidence that either the word frequency effect or the neighborhood size effect was modulated by print exposure.

Response Latencies and Error Rates for Nonwords. The mean response latencies and error rates to the nonword stimuli are shown in Table 3. As expected, responses to nonwords with large neighborhoods were slower than the responses to nonwords with small neighborhoods (628 ms vs. 606 ms), $F_s(1, 73) = 51.39$, $p < .001$, MSE = 347.03; $F_i(1, 260) = 13.25$, $p < .001$, MSE = 2,305.12, and more errors were made to nonwords with large neighborhoods than to nonwords with small neighborhoods (6.7% vs. 3.6%), $F_s(1, 73) = 43.37$, $p < .001$, MSE = 7.68; $F_i(1, 260) = 15.60$, $p < .001$, MSE = 37.59. The main effect of group was not significant in the subject analysis of response latencies ($F_s < 1$), but the effect was significant in the item analysis, $F_i(1, 260) = 7.25$, $p < .01$, MSE = 2,305.12. The main effect of group was not significant in the analysis of errors, $F_s(1, 73) = 1.92$, $p > .15$, MSE = 26.84; $F_i(1, 260) = 2.41$, $p > .10$, MSE = 37.59, and group did not interact with neighborhood size in the response latency analysis, $F_s(1, 73) = 2.32$, $p > .10$, MSE = 347.03; $F_i < 1$, nor in the error analysis, $F_s(1, 73) = 3.34$, $p = .07$, MSE = 7.68; $F_i(1, 260) = 1.20$, $p > .25$, MSE = 37.59. As can be seen in Table 3, apart from the slightly slower responses of the low ART participants, the two groups' responses to the nonword stimuli were very similar. In particular, the difference between the nonwords with large neighborhoods and the nonwords with small neighborhoods (the inhibitory neighborhood size effect) was quite similar for the two groups of participants (i.e., 26 ms for the low ART group and 18 ms for the high ART group). The low ART participants thus had no more difficulty rejecting very word-like nonwords than the high ART participants.

Discussion

The results of this experiment nicely replicate those of Lewellen et al. (1993), in that the neighborhood size effect and the word frequency effect were not modulated by print exposure. In addition, print exposure appeared to have little effect (if any) on the responding to nonword stimuli. These results are very different from the results of Chateau and Jared (2000), who reported a larger word frequency effect and a larger neighborhood size effect for participants with low levels of print exposure.

Experiment 2

One important difference between Lewellen et al.'s (1993) experiment and Chateau and Jared's (2000) experiment was the type of nonword stimuli used in the lexical decision task. The nonwords used in the Chateau and Jared's experiment were pseudohomophones (e.g., BRANE), whereas the nonwords used in Lewellen et al.'s experiment (and in Experiment 1) were orthographically-legal and pronounceable, but were not pseudohomophones (e.g., BRINT). The possibility that this difference accounts for the very different outcomes of these experiments must now be explored. To this end, in Experiment 2 the same word stimuli were presented to a new group of participants (again divided into high and low ART groups) and a set of pseudohomophones was used as nonword stimuli.

Method

Participants. One-hundred and twenty University of Calgary undergraduate students participated in this experiment. All participants received bonus credit toward a psychology course in exchange for their participation. The participants' ages ranged from seventeen to forty-nine years of age. All participants were native English speakers and reported normal or corrected to normal vision.

Participants were administered the ART at the start of the experimental session. For the data analyses, two groups of participants were formed using a tertile split of the ART scores. Thirty-eight participants were assigned to the high ART group (with a mean ART score of 25.9) and thirty-eight were assigned to the low ART group (with a mean ART score of 8.0). Note that the mean ART scores of the high and the low ART groups in this experiment were very similar to the mean ART scores of the high and the low ART groups in Experiment 1 (26.4 and 10.0, respectively).

Stimuli. The word stimuli were identical to those used in Experiment 1. One-hundred and thirty-two pseudohomophones were presented in the experiment. The pseudohomophones were matched to the words on length, such that 72 were four letters in length and 60 were five letters in length. Forty of these stimuli were used in Chateau and Jared's (2000) study; twenty were "common" pseudohomophones (e.g., REECH) and twenty were "unusual" pseudohomophones (e.g., CAWPH). The remaining 92 pseudohomophones consisted of stimuli used in other studies (e.g., Stone & Van Orden, 1993). All of these stimuli were "common" pseudohomophones according to Chateau and Jared's definition.

Apparatus and Procedure. The apparatus and procedure were identical to those of Experiment 1. Each participant completed 16 practice trials prior to the collection of data. The practice stimuli consisted of eight words (four of low-frequency, four of high frequency) and eight pseudohomophones. These practice stimuli were not used in the experiment, and the data from these practice trials were not analyzed. The procedure for these practice trials was identical to that of Experiment 1.

Results

For the word data, response latencies and error rates were submitted to a 2 (Group: high ART, low ART) x 2 (Word Frequency: high, low) x 2 (Neighborhood Size: small, large) mixed-model factorial ANOVA. For the nonword data, response latencies and error rates were submitted to a 2 (Group: high ART, low ART) x 2 (Pseudohomophone Type: common, unusual) mixed-model ANOVA. Both subject (F_s) and item (F_i) analyses were carried out. Response latencies to words or to pseudohomophones that were less than 250 ms or greater than 1500 ms were considered outliers and were removed from the data set. A total of 171 observations (0.85% of the data) were removed by this procedure (97 response latencies of the high ART participants and 74 response latencies of the low ART participants).

Response Latencies for Words. The mean response latencies of correct responses and the mean error rates are shown in Table 4. The main effect of group was significant in the item analysis, $F_i(1, 256) = 14.15$, $p < .001$, $MSE = 1{,}475.62$, but not in the subject analysis, $F_s < 1$. Overall, the response latencies of the low ART participants were slightly slower than the response latencies of the high ART participants (557 ms vs. 542 ms). The main effect of word frequency was significant, $F_s(1, 74) = 300.96$, $p < .001$, $MSE = 436.81$; $F_i(1, 256) = 81.95$, $p < .001$, $MSE = 1{,}475.62$, as was the main effect of neighborhood size, $F_s(1, 74) = 53.95$, $p < .001$, $MSE = 382.88$; $F_i(1, 256) = 13.24$, $p < .001$, $MSE = 1{,}475.62$. Responses to high-frequency words were an average of 41 ms faster than responses to low-frequency words, and responses to words with large neighborhoods were an average of 17 ms faster than responses to words with small neighborhoods. Again, as expected, the interaction between word frequency and neighborhood size was significant, $F_s(1, 74) = 33.71$, $p < .001$, $MSE = 359.17$; $F_i(1, 256) = 7.16$, $p < .01$, $MSE = 1{,}475.62$. For the low-frequency words, words with large neighborhoods were responded to an average of 29 ms faster than words with small neighborhoods, $F_s(1, 74) = 67.66$, $p < .001$, $MSE = 475.88$; $F_i(1, 128) = 14.98$, $p < .001$, $MSE = 1{,}982.45$. For the high-frequency words the neighborhood size effect was only 4 ms (both p's $> .10$).

Unlike Experiment 1, there were significant interactions between group and word frequency, $F_s(1, 74) = 19.62$, $p < .001$, $MSE = 436.81$; $F_i(1, 256) = 6.67$, $p < .05$, $MSE = 1{,}475.62$, and between group and neighborhood size, $F_s(1, 74) = 4.57$, $p < .05$, $MSE = 382.88$; $F_i(1, 256) = 1.55$, $p > .20$, $MSE = 1{,}475.62$. The three-way interaction between group, word frequency, and neighborhood size was also significant, $F_s(1, 74) = 6.16$, $p < .05$, $MSE = 359.17$; $F_i(1, 256) = 1.31$, $p > .20$, $MSE = 1{,}475.62$. An examination of Table 4 reveals the source of these interactions. The word frequency effect was larger for the low ART participants than for the high ART participants (52 ms vs. 31 ms), and the neighborhood size effect was larger for the low ART participants than for the high ART participants (21 ms vs. 12 ms). The latter difference was, not surprisingly, confined to the low-frequency words, hence the three-way interaction. For the low ART group, the neighborhood size effect for low-frequency words was 40 ms, and for the high ART group it was 19 ms, whereas for the high-frequency words the effects were 3 ms and 5 ms, respectively.

Error Rates for Words. In the analysis of error rates there were main effects of group, $F_s(1, 74) = 29.77$, $p < .001$, $MSE = 27.63$; $F_i(1, 256) = 22.85$, $p < .001$, $MSE = 31.27$, word frequency, $F_s(1, 74) = 68.31$, $p < .001$, $MSE = 15.80$; $F_i(1, 256) = 29.98$, $p < .001$, $MSE = 31.27$, and neighborhood size, $F_s(1, 74) = 11.52$, $p < .01$, $MSE = 13.98$; $F_i(1, 256) = 4.47$, $p < .05$, $MSE = 31.27$. Consistent with the response latency analysis, the low ART participants made more errors than the high ART participants (5.6% vs. 2.3%), fewer errors were made to high-frequency words than to low-frequency words (2.1% vs. 5.8%), and fewer errors were made to words with large neighborhoods than to words with small neighborhoods (3.2% vs. 4.7%). Also consistent with the response latency analysis was the interaction between word frequency and neighborhood size, $F_s(1, 74) = 9.15$, $p < .01$, MSE = 13.95; $F_i(1, 256) = 3.55$, $p = .06$, MSE = 31.27.

There was a significant interaction between group and word frequency, $F_s(1, 74) = 17.99$, $p < .001$, MSE = 15.80; $F_i(1, 256) = 7.90$, $p < .01$, MSE = 31.27, which mirrored the results obtained in the response latency analysis. In terms of error rates, the word frequency effect was larger for the low ART group, as the difference between the high- and the low-frequency words was larger for the low ART participants (5.7%) than for the high ART participants (1.8%). Unlike the response latency analysis, the interaction between group and neighborhood size was not significant, $F_s(1, 74) = 2.15$, $p > .10$, MSE = 13.98; $F_i < 1$, nor was the interaction between group, word frequency, and neighborhood size (both F's < 1). As can be seen in Table 4, however, the pattern of error rates is entirely consistent with the pattern of response latencies. In particular, the neighborhood size effect for low-frequency words was larger for the low ART participants than for the high ART participants (3.8% vs. 1.7%, respectively), and neighborhood size had no effect on the error rates to high-frequency words for either group of participants.

Response Latencies and Error Rates for Pseudohomophones. Recall that in Chateau and Jared's (2000) study, for common pseudohomophones (e.g., REECH), low ART participants were slower and more error prone than high ART participants, but for unusual pseudohomophones (e.g., CAWPH) there were no group differences. This produced interactions between group and pseudohomophone type in Chateau and Jared's response latency and error analyses.

As can be seen in Table 3, Chateau and Jared's stimuli produced the same pattern of effects in this experiment as well. In the analysis of response latencies, the main effect of group was significant in the item analysis, $F_i(1, 76) = 15.72$, $p < .001$, $MSE = 1{,}382.82$, but not in the subject analysis, $F_s(1, 74) = 1.76$, $p > .15$, $MSE = 2{,}0696.64$ (presumably due to the lower statistical power). More importantly, the main effect of pseudohomophone type (common vs. unusual) was significant, $F_s(1, 74) = 222.10$, $p < .001$, $MSE = 1{,}336.65$; $F_i(1, 76) = 107.39$, $p < .001$, $MSE = 1{,}382.82$, and so was the interaction between group and pseudohomophone type, $F_s(1, 74) = 8.68$, $p < .01$, $MSE = 1{,}336.65$; $F_i(1, 76) = 4.75$, $p < .05$, $MSE = 1{,}382.82$. Similarly, in the analysis of errors, the main effect of group was significant, $F_s(1, 74) = 4.19$, $p < .05$, $MSE = 30.76$; $F_i(1, 76) = 2.45$, $p > .10$, $MSE = 27.66$, the main effect of pseudohomophone type was significant, $F_s(1, 74) = 65.47$, $p < .001$, $MSE = 29.30$; $F_i(1, 76) = 36.50$, $p < .001$, $MSE = 27.66$, and so was the interaction, $F_s(1, 74) = 8.98$, $p < .01$, $MSE = 29.30$; $F_i(1, 76) = 5.01$, $p < .05$, $MSE = 27.66$. As Chateau and Jared reported, for the common pseudohomophones, the low ART participants were slower and more error prone than the high ART participants, $F_s(1, 74) = 3.72$, $p = .05$, $MSE = 11{,}992.53$; $F_i(1, 38) = 13.25$, $p < .01$, $MSE = 1968.67$, and $F_s(1, 74) = 6.64$, $p < .05$, $MSE = 57.29$; $F_i(1, 38) = 3.72$, $p = .06$, $MSE = 53.76$, respectively. For the unusual pseudohomophones, the two groups did not differ in their response latencies (both p's > .10), and the low ART participants actually made slightly fewer errors (0.2%) than the high ART participants (1.0%), $F_s(1, 74) = 4.27$, $p < .05$, $MSE = 2.77$; $F_i(1, 38) = 3.98$, $p = .05$, $MSE = 1.57$.

Also listed in Table 3 are the response latencies and error rates to all the "common" pseudohomophones presented in the experiment (i.e., the 20 "common" pseudohomophones from the Chateau and Jared study, and the additional 92 "common" pseudohomophones presented in this experiment). An analysis of this larger set of pseudohomophones led to the same conclusions. Specifically, the low ART participants were slower to reject the pseudohomophones than the high ART participants, $F_s(1, 74) = 3.23$, $p = .07$, MSE = 11,365.18; $F_i(1, 222) = 50.67$, $p < .001$, MSE = 2,479.07, and also committed more errors, $F_s(1, 74) = 10.72$, $p < .01$, MSE = 32.06; $F_i(1, 222) = 15.94$, $p < .001$, MSE = 63.52.

Finally, the data were also examined for evidence of a baseword frequency effect, particularly one that interacted with print exposure. A number of investigators have reported that responses to pseudohomophones based on high-frequency words (e.g., MAIK) are faster and more accurate than responses to pseudohomophones based on low-frequency words (e.g., GLEW). Most of these studies have used the pronunciation task (for a review, see Borowsky & Masson, 1999), but there are reports of baseword frequency effects in lexical decision tasks (Van Orden, 1991; Van Orden et al, 1992; although see McCann, Besner, & Davelaar, 1988, and Seidenberg, Petersen, MacDonald, & Plaut, 1996, for unsuccessful attempts). The baseword frequency effect suggests that pseudohomophones, like words, activate frequency-sensitive word representation in the mental lexicon. If so, then one might expect that the baseword frequency effect, like the standard word frequency effect, would vary as a function of print exposure.

For the purposes of this analysis, the 112 "common" pseudohomophones presented in the experiment were categorized according to the frequency of their base word. Pseudohomophones based on words with normative frequencies greater than or equal to 100 were considered to have a high-frequency baseword ($N = 27$), and pseudohomophones based on words with normative frequencies less than or equal to 50 were considered to have a low-frequency baseword ($N = 63$). Pseudohomophones based on words with frequencies between 51 and 99 (inclusive) were not used in this analysis. The response latency and error data were then submitted to a 2 (Group: low ART, high ART) x 2 (Baseword Frequency: low, high) ANOVA.

In the analysis of response latencies, there was no baseword frequency effect ($F_i < 1$), nor was there an interaction between group and baseword frequency ($F_i < 1$). For the low ART group, the pseudohomophones based on high-frequency words were responded to no faster than the pseudohomophones based on low-frequency words (685 ms vs. 681 ms, respectively). The same was true for the high ART participants, with mean response latencies of 637 ms for the pseudohomophones based on high-frequency words and 631 ms for pseudohomophones based on low-frequency words. In the analysis of error rates, the baseword frequency effect was statistically significant, $F_i(1, 176) = 6.45$, $p < .05$, $MSE = 68.72$. Pseudohomophones based on high-frequency words were responded to more accurately (i.e., correctly rejected more often) than pseudohomophones based on low-frequency words (5.9% vs. 9.4%, respectively). But again, there was no interaction between group and baseword frequency ($F_i < 1$). Thus, although there was some evidence of a baseword frequency effect in this experiment, there was no indication that it interacted with print exposure.

Combined Analyses of Word Data. Combined analyses of the word data from Experiments 1 and 2 revealed several major differences between the two sets of results. First, considering only the low ART participants, participants' responses were generally slower and more error prone in Experiment 2 than in Experiment 1, as indicated by a significant effect of experiment in a combined analysis of response latencies, $F_s(1, 73) = 1.73$, $p > .15$, $MSE = 19451.54$; $F_i(1, 256) = 18.52$, $p < .001$, $MSE = 1{,}748.60$, and in a combined analysis of errors, $F_s(1, 73) = 6.44$, $p < .05$, $MSE = 34.04$; $F_i(1, 256) = 4.96$, $p < .05$, $MSE = 38.89$. Second, the word frequency effect was slightly larger in Experiment 2 than in Experiment 1, as suggested by a marginally significant interaction between experiment and word frequency in the response latency analysis (53 ms vs. 43 ms), $F_s(1, 73) = 3.44$, $p = .06$, $MSE = 489.36$; $F_i(1, 256) = 1.25$, p

> .20, MSE = 1,748.60, and a significant interaction in the error analysis (5.7% vs. 3.0%), $F_s(1, 73) = 6.86$, $p < .05$, MSE = 18.33; $F_i(1, 256) = 2.85$, $p = .09$, MSE = 38.89. Third, the neighborhood size effect for low-frequency words was larger in Experiment 2 (40 ms) than in Experiment 1 (20 ms), which produced a significant interaction between experiment, word frequency, and neighborhood size, $F_s(1, 73) = 4.27$, $p < .05$, MSE = 349.28; $F_i < 1$, and a significant interaction between experiment and neighborhood size in a combined analysis of only low-frequency words, $F_s(1, 73) = 6.18$, $p < .05$, MSE = 520.64; $F_i(1, 128) = 1.33$, $p > .20$, MSE = 2496.47. Together these results indicate that, for low ART participants, pseudohomophones slowed the responses to words and produced larger word frequency and neighborhood size effects relative to when regular nonwords were used.

In contrast, the high ART participants' data in Experiment 1 was almost identical to the high ART participants' data in Experiment 2. There was no effect of experiment in a combined analysis of response latencies (both F's < 1) or of errors, $F_s(1, 74) = 2.14$, $p > .10$, MSE = 14.43; $F_i(1, 256) = 2.51$, $p > .10$, MSE = 10.70, as the response latencies and error rates of the high ART participants in the two experiments were essentially the same. In addition, there were no interactions between experiment and word frequency or between experiment and neighborhood size in either the response latency analysis or in the error analysis (all p's > .15). The absence of significant interactions is not surprising given that the word frequency effect was 35 ms in Experiment 1 and 31 ms in Experiment 2, and that the neighborhood size effect for low-frequency words was 17 ms in Experiment 1 and 19 ms in Experiment 2.

A combined analysis of all the word data from the two experiments (Experiment x Group x Word Frequency x Neighborhood Size) produced several interactions that supported the conclusions reached above. First, the interaction between experiment and group was significant in the item analysis of response latencies, $\underline{F}_i(1, 512) = 9.62$, $\underline{p} > .01$, $\underline{MSE} = 1{,}457.40$, although not in the subject analysis ($\underline{F}_s < 1$). In the error analysis, this interaction was significant in both analyses, $\underline{F}_s(1, 147) = 8.61$, $\underline{p} < .01$, $\underline{MSE} = 24.17$; $\underline{F}_i(1, 512) = 7.34$, $\underline{p} < .01$, $\underline{MSE} = 24.80$. These interactions reflected the fact that, as already noted, the low ART participants in Experiment 2 had slower response latencies and higher error rates than the low ART participants in Experiment 1, whereas the response latencies and error rates of the high ART participants in the two experiments was statistically equivalent. The significant interaction between experiment, group, and word frequency in the response latency analysis, $\underline{F}_s(1, 147) = 4.34$, $\underline{p} < .05$, $\underline{MSE} = 441.30$; $\underline{F}_i(1, 512) = 1.85$, $\underline{p} > .15$, $\underline{MSE} = 1457.40$, and in the error analysis, $\underline{F}_s(1, 147) = 8.54$, $\underline{p} < .01$, $\underline{MSE} = 13.90$; $\underline{F}_i(1, 512) = 4.18$, $\underline{p} < .05$, $\underline{MSE} = 24.80$, was further evidence that the word frequency effect was larger in Experiment 2 than in Experiment 1, but only for the low ART participants.

Combined Analyses of Nonword Data. Finally, the nonword data from Experiments 1 and 2 were compared to determine if the pseudohomophones used in Experiment 2 were more difficult to reject than the regular nonwords used in Experiment 1. Of particular interest were comparisons between the pseudohomophones and the large neighborhood nonwords used in Experiment 1 (see Table 3). Like the "common" pseudohomophones, these nonwords were pronounceable and very word-like (e.g., WINT), but when pronounced they did not sound like words. Consequently, by comparing the responses to the large neighborhood nonwords to the pseudohomophones one can assess any additional difficulty participants experienced rejecting nonwords because they sounded like words. It is well-known that lexical decision responses to pseudohomophones are slower and more error prone than responses to matched nonword controls that do not sound like words (e.g., JALE vs. JARL; Coltheart et al., 1977; Vanhoy &

Van Orden, 2001; for reviews see Berent & Perfetti, 1995; Frost, 1998; Van Orden, Pennington, Stone, 1990). Of interest here is the possibility that this so-called pseudohomophone effect interacts with print exposure.

As can be seen in Table 3, for the low ART participants, responses to pseudohomophones were slower and more error prone than responses to large neighborhood nonwords (679 vs. 638 ms, and 10.1% vs. 7.7%). The 41 ms latency difference was marginally significant in a subject analysis, $F_s(1, 73) = 3.09$, $p = .08$, MSE = 10097.30, and was statistically significant in an item analysis, $F_i(1, 176) = 24.19$, $p < .001$, MSE = 2758.81. The error difference was marginally significant in both analyses, $F_s(1, 73) = 3.47$, $p = .06$, MSE = 29.73; $F_i(1, 176) = 2.75$, $p = .09$, MSE = 83.15. These results suggest that the low ART participants had more difficulty rejecting pseudohomophones than regular nonwords. An examination of Table 3 reveals that these differences were not as pronounced for the high ART participants. In particular, for the high ART participants responses to pseudohomophones were only 16 ms slower than responses to regular nonwords, $F_s < 1$; $F_i(1, 176) = 4.43$, $p < 05$, MSE = 1889.94, and the error rates to these stimuli were virtually identical (5.7% vs. 5.8%), $F_s < 1$; $F_i < 1$. Together these results suggest that the high ART participants did not experience as much difficulty rejecting the pseudohomophones as the low ART participants. Because the only important difference between the large neighborhood nonwords and the pseudohomophones was whether the stimulus sounded like a word (e.g., BRANE vs. BRINT), the implication is that the two groups differed in the extent to which they were able to suppress phonological activation from the pseudohomophones or, alternatively, the extent to which they relied on phonological activation to make their lexical decision responses. These issues will be discussed further in the General Discussion.

Discussion

The results of this experiment replicate those of Chateau and Jared (2000), who found that the word frequency effect and the neighborhood size effect were modulated by print exposure. Given that the identical word stimuli were used in Experiments 1 and 2, it would seem reasonable to conclude that it was the different type of nonword stimuli that was responsible for the different outcomes observed in these experiments. That is, taken together, the results of Experiments 1 and 2 suggest that when pseudohomophones are used in a lexical decision task one can observe the effects of print exposure, but when regular nonwords are used one cannot. To be sure this is the case, in Experiment 3 another attempt was made to observe an effect of print exposure when regular nonwords are used in the lexical decision task.

Experiment 3

Like Experiment 1, the purpose of this experiment was to determine if the word frequency effect and the neighborhood size effect would be modulated by print exposure when regular nonwords were used in a lexical decision task. To increase the possibility of observing these effects, and to increase the generalizability of any conclusions, a very large set of word stimuli was used. The stimulus sets that have been used in the studies thus far have been fairly small. In Chateau and Jared's (2000) experiment, for example, there were 48 words used in the lexical decision task. In Experiment 1 of the present study, 140 words were used (72 four-letter words and 60 five-letter words), but given that there are approximately 1,500 four-letter words and 2,100 five-letter words listed in the Kucera and Francis (1967) norms, clearly only a small fraction of the population of words has been sampled. In this experiment, the responses to 300 of the 519 three-letter words listed in the Kucera and Francis norms were collected (57.8%).

The size of the data set also permitted an examination of other orthographic variables whose effect could be modulated by print exposure. Other than printed word frequency and neighborhood size, no other variables have been examined in these type of studies. One variable of interest is length-specific positional bigram frequency (i.e., the frequency with which a particular pair of letters occur in a specified position of words of a given length; Mayzner & Tresselt, 1965). Whether or not there is a bigram frequency effect in the lexical decision task is not at all clear (see Gernsbacher, 1984, for a review), but the possibility that this orthographic redundancy effect varies as a function of print exposure has never been examined.

Another variable of interest is the frequency of a word's neighbors, referred to as neighborhood frequency. The neighborhood size manipulations in previous experiments including Experiments 1 and 2 did not distinguish between neighbors that were lower in frequency than the word itself and neighbors that were higher in frequency. A number of investigators have examined the impact on lexical decision response latencies of having higher frequency neighbors (e.g., Carreiras, Perea, & Grainger, 1997; Forster & Shen, 1996; Grainger, 1990; Grainger & Jacobs, 1996; Grainger, O'Regan, Jacobs, & Segui, 1989, 1992; Grainger & Segui, 1990; Huntsman & Lima, 1996; Perea & Pollatsek, 1998; Sears et al., 1995).

Most of these studies seem to show that lexical decision latencies to low-frequency words with higher frequency neighbors are slower than those to low-frequency words without higher frequency neighbors (usually referred to as an "inhibitory neighborhood frequency effect"). But the effect of neighborhood frequency is less consistent when English stimuli are used. That is, some investigators have reported inhibitory neighborhood frequency effects (Huntsman & Lima, 1996; Perea & Pollatsek, 1998), whereas other investigators have reported either null or facilitatory neighborhood frequency effects (Forster & Shen, 1996; Paap and Johansen, 1994; Sears et al., 1995; Siakaluk et al., 2001).

Paap and Johansen (1994) reported that, in a regression analysis of lexical decision latencies to words, the number of lower frequency neighbors and the number of higher frequency neighbors both accounted for a significant percentage of variance. According to their analysis, an increase in the number of lower frequency neighbors was associated with a decrease in the lexical decision latency, whereas an increase in the number of higher frequency neighbors was associated with an increase in lexical decision latency. The data from Experiment 3 were submitted to similar regression analyses, and the effect of print exposure on the neighborhood frequency effect was assesed via semipartial correlation coefficients.

Method

Participants. One-hundred University of Calgary undergraduate students participated in the experiment. All participants received bonus credit toward a psychology course in exchange for their participation. The mean age of the participants was 22.1 years, with a range of seventeen to forty years of age. All were native English speakers and reported normal or corrected to normal vision.

The participants were administered the ART and the Comprehension subtest of the Nelson-Denny Reading Test during the experimental session. For the data analyses, two groups of participants were formed using a tertile split of the ART scores. Twenty-six participants were assigned to the high ART group (M = 26.0) and twenty-eight participants were assigned to the low ART group (M = 9.8). Note that the mean ART scores of the high and low ART groups in this experiment were very similar to the mean ART scores of the high and low ART groups in Experiment 1 (26.4 and 10.0, respectively) and in Experiment 2 (25.9 and 8.0, respectively). Not surprisingly, the mean Nelson-Denny comprehension score of the low ART group (25.0) was lower than the mean score of the high ART group (32.1), $t(52) = 4.57$, $p < .001$. The correlation between the ART scores and the Nelson-Denny scores was .53 ($p < .001$).

Stimuli. All of the stimuli were three letters in length. Three-hundred words and 300 orthographically-legal and pronounceable nonwords (e.g., BAP) were presented in the experiment. For each word the Kucera and Francis (1967) normative frequency, the number of lower frequency neighbors, the number of higher frequency neighbors, and the summed positional bigram frequency was determined.

The subjective frequency of each word was also determined to provide an alternative measure of word frequency, given that the Kucera and Francis (1967) printed frequency norms tend to be somewhat unreliable for low-frequency words (Gernsbacher, 1984; Gordon, 1985). In a separate study, 204 undergraduate students were asked to estimate how frequently they encountered 444 different words in print, using a scale from 0 (Very Infrequently) to 9 (Very Frequently). The words were three, four, and five letters in length, and were listed in a random order on five sheets of paper. Three-hundred of these words were used in Experiment 3.

The descriptive statistics (mean and range) for the word stimuli were as follows: Kucera and Francis (1967) normative frequency (99.0, 1-392); subjective frequency (4.14, 1-8); number of lower frequency neighbors (6.1, 0-20); number of higher frequency neighbors (5.43, 0-20); summed positional bigram frequency (1074.9, 2-15,105). The mean neighborhood size of the words was 11.5; for the nonwords the mean neighborhood size was 12.8.

Apparatus and Procedure. The apparatus was identical to that used in Experiments 1 and 2. The stimuli were presented in two blocks of 300 items (150 words and 150 nonwords in each block). Participants completed the Comprehension subtest of the Nelson-Denny between the two blocks. The order in which the blocks were presented was counterbalanced across participants. Each block began with 16 practice trials, which consisted of 8 three-letter words and 8 three-letter nonwords. (These practice stimuli were not used in the experiment, and the data from these practice trials were not analyzed.) The procedure for these practice trials was identical to that of the Experiments 1 and 2.

Results

Prior to the multiple regression analyses, the mean response latency and mean error rate for each word was calculated, and words with an error rate greater than 50% were excluded from the multiple regression analyses. Ten words were excluded in this fashion. A logarithmic transformation of Kucera and Francis's (1967) normative frequencies was employed in the regression analyses. Preliminary analyses revealed that the correlation between log normative frequency and subjective frequency was unacceptably high (.80), and so separate regression analyses were conducted with these two variables.

High ART Group. In the first set of analyses, the predictor variables were log normative frequency, the number of lower frequency neighbors, the number of higher frequency neighbors, and bigram frequency. The predictor variables were entered simultaneously in all the analyses. Semi-partial correlation coefficients were computed to assess the unique correlation between the mean word latencies and error rates and each of the predictor variables; these are listed in Table 5. Together these variables explained 18.1% of the variance in the word latencies, $F(4, 285) = 15.75$, $p < .001$, $MSE = 3,365.98$, but log word frequency was the only significant predictor. In the analysis of error rates, 6.8% of the variance was accounted for, $F(4, 285) = 5.24$, $p < .001$, $MSE = 68.91$, but no single predictor accounted for a significant portion of the variance. When subjective frequency ratings were used as predictor (instead of log normative frequency), 29.7%

of the variance in the word latencies was accounted for, $F(4, 285) = 30.15$, $p < .001$, MSE = 2, 887.92, and 11.8% of the variability in errors was accounted for, $F(4, 285) = 9.62$, $p < .001$, MSE = 65.18, but subjective frequency was the only significant predictor. These data are listed in Table 6.

Because most investigators have focused on orthographic neighborhood effects for low frequency words, separate regression analyses were conducted on this subset of the stimuli (i.e., words with normative frequencies less than or equal to 50; N = 208). In the analysis using log word frequency (Table 5), 17.6% of the variance in the word latencies was accounted for, $F(4, 203) = 10.85$, $p < .001$, MSE = 3,806.56, with log word frequency the only significant predictor. In the analysis of errors, 11.8% of the variance was accounted for, $F(4, 203) = 6.79$, $p < .001$, MSE = 76.60, and the number of lower frequency neighbors was the only predictor which approached statistical significance ($p = .06$).

The analyses using subjective frequency ratings produced very different results (Table 6). Specifically, in the analysis of low-frequency word latencies the predictors accounted for 38.5% of the variance, $F(4, 203) = 31.87$, $p < .001$, MSE = 2,838.19, with subjective frequency and the number of lower frequency neighbors being the only significant predictors. Subjective frequency and the number of lower frequency neighbors were also the only significant predictors in the error analysis, which explained 23.9% of the variance, $F(4, 203) = 15.93$, $p < .001$, MSE = 66.10. Thus, the analyses using subjective frequency accounted for a much larger percentage of the variance in word latencies and in errors than the analyses using log word frequency (38.5% vs. 17.6%, and 23.9% vs. 11.8%, respectively), and only in the former case did the number of lower frequency neighbors account for any additional variance.

Low ART Group. Tables 5 and 6 also list the semi-partial correlation coefficients from the regression analyses of the low ART participants' data. In the analysis of word latencies, using log word frequency, 22.0% of the variance was accounted for, $F(4, 285) = 20.16$, $p < .001$, MSE = 4,941.06, with significant semi-partial correlations for log word frequency and the number of lower frequency neighbors. In the analysis of errors, only log word frequency was a significant predictor, with 13.4% of the variance explained by this variable, $F(4, 285) = 11.02$, $p < .001$, MSE = 144.96. Using subjective frequency ratings, 35.6% of the variance in word latencies was explained, $F(4, 285) = 39.51$, $p < .001$, MSE = 4077.93, and the statistically significant predictors were subjective frequency and the number of lower frequency neighbors. In the analysis of word errors, 24.9% of the variance was explained, $F(4, 285) = 23.73$, $p < .001$, MSE = 125.56, but subjective frequency was the only significant predictor.

As can be seen in Table 5, separate analyses of the low-frequency words produced similar results. In the response latency analysis using log word frequency, 17.5% of the variance was accounted for, $F(4, 203) = 10.77$, $p < .001$, MSE = 5577.08, and log word frequency and the number of lower frequency neighbors were significant predictors. In the error analysis, log word frequency was the only significant predictor, with 14.4% of the variance in word errors accounted for, $F(4, 203) = 8.58$, $p < .001$, MSE = 172.55. In the analyses using subjective frequency ratings, 39.1% of the variance in word latencies was explained, $F(4, 203) = 32.67$, $p < .001$, MSE = 4112.95, and 35.7% of the variance in word errors was explained, $F(4, 203) = 28.20$, $p < .001$, MSE = 129.66. Subjective frequency and the number of lower frequency neighbors were the only significant predictors in these analyses. Note that, as was the case in the analyses of the High ART group's data, the analyses using subjective frequency accounted for a much larger percentage of the variance in word latencies and in errors than the analyses using log word frequency (39.1% vs. 17.5%, and 35.7% vs. 14.4%, respectively).

Discussion

Like Experiment 1, the purpose of this experiment was to determine if the word frequency effect and the neighborhood size effect would be modulated by print exposure when regular nonwords were used in a lexical decision task. To increase the possibility of observing such an effect, a very large set of word stimuli was used. In addition, the possibility that other orthographic variables such as, subjective frequency, the number of higher frequency neighbors, and bigram frequency could be modulated by print exposure was also examined.

The results of this experiment are easily summarized. First, the only significant predictors of lexical decision performance were word frequency (i.e., log normative frequency and subjective frequency) and the number of lower frequency neighbors. The number of higher frequency neighbors and bigram frequency did not account for a significant percentage of variance in any of the regression analyses. Second, there was no indication that the word frequency effect or the neighborhood size effect was modulated by print exposure. The semipartial correlation coefficients for the word frequency effect were almost identical in every analysis of the low and high ART group's data, and the semi-partial correlations for the number of lower frequency neighbors were of very similar magnitude. In the regression analysis that explained the most variance in the high and low print exposure group's data (the analysis of only the low-frequency words, using subjective frequency as a predictor), the semipartial correlations for subjective frequency were identical (.49), and the semipartial correlations for the number of lower frequency neighbors were statistically equivalent (.21 vs. .16). These results lend strong support to the notion that, in a lexical decision task, print exposure modulates the effect of word frequency and the effect of neighborhood size only when pseudohomophones are used as distractor stimuli.

General Discussion

The purpose of this study was to re-examine whether print exposure modulates the effect of word frequency and the effect of neighborhood size, given the conflicting finds of Lewellen et al. (1993) and Chateau and Jared (2000). In Experiment 1, like the Lewellen et al. experiment, regular nonwords were used in the lexical task. There was no evidence that print exposure had any effect on the responses to words or to nonwords. This result was confirmed in Experiment 3.

In Experiment 2, like the Chateau and Jared experiment, pseudohomophones were used in the lexical decision task. In this experiment there was clear evidence of an effect of print exposure on the responses to words and to pseudohomophones. In particular, participants with lower levels of print exposure showed larger word frequency effects and larger neighborhood size effects than participants with higher levels of print exposure, and they were also slower to reject pseudohomophones. Several analyses revealed that the low print exposure participants were slower and more error prone in Experiment 2 than in Experiment 1, and, in addition, exhibited a larger word frequency effect and a larger neighborhood size effect in Experiment 2. Their responses to the pseudohomophones used in Experiment 2 were also slower and more error prone than their responses to the regular nonwords used in Experiment 1. Overall, there was unmistakable evidence that the low print exposure participants responded differently than both the high print exposure participants in Experiment 2 and the low print exposure participants in Experiment 1. In contrast, the responses of the high print exposure participants in both experiments were virtually identical. For these participants, there was little evidence that even the responses to pseudohomophones were any slower than the responses to regular nonwords.

These results pose two interrelated questions. First, why is there an effect of print exposure on the word frequency effect and the neighborhood size effect only when pseudohomophones are used in the lexical decision task? And second, why are the responses of high print exposure participants virtually unaffected by nonword manipulations (i.e., regular nonwords vs. pseudohomophones)?

Stone and Van Orden (1998) demonstrated that the type of nonwords used in a lexical decision task has a very large impact on the magnitude of the word frequency effect. In their experiments, when the nonwords were orthographically illegal letter strings (e.g., BTESE) the word frequency effect was 36 ms, when the nonwords were orthographically legal (e.g., DEEST) the frequency effect was 76 ms, and when the nonwords were pseudohomophones (e.g., BEEST) the frequency effect was 159 ms. Neighborhood size was not manipulated in Stone and Van Orden's study, but Siakaluk et al. (2001) reported that both the word frequency effect and the neighborhood size effect are when the nonwords have many orthographic neighbors (e.g., CLOW) than when they have none (e.g., BISM). Consequently, it is reasonable to conclude that, like the word frequency effect, the magnitude of the neighborhood size effect increases when the nonwords are more word-like (i.e., when they have a large neighborhood or when they sound like words). In this respect, for the low ART participants, the increase in the word frequency effect and the neighborhood size effect between Experiments 1 and 2 makes a great deal of sense. As the task becomes more difficult, due to the increased difficulty of distinguishing the words from the nonwords, orthographic variables such as word frequency and neighborhood size have a larger impact on performance. What this does not explain, of course, is why the high ART participants were not affected in the same manner.

It may be that high print exposure participants use different word recognition processes than low print exposure participants. Of particular relevance is the research that has sought to determine whether readers use phonological information, orthographic information, or both to access the meaning of words from print (e.g., Coltheart, et al., 1977; Doctor & Coltheart, 1978; Frost, 1998; Van Orden, 1987). One possibility is that readers use their knowledge of the correspondences between letters and sounds to translate printed letters into their phonological representations and use the sound information to activate word meanings (the phonological route). The second possibility is that readers obtain meaning directly from the printed word, ignoring the letters that encode phonology (the direct route). This is one of the most studied issues in reading research. The empirical question is which of these processes is actually used by high print exposure participants and which one is used by low print exposure participants?

There is considerable evidence that indicates that the phonological route is critical when children are learning to read (Elboro, 1996). For example, Elboro argues that phonemic awareness (the ability to detect individual phonemes in spoken words) is a cornerstone in the development of reading skills and that it would be difficult to explain the role phonemic awareness if skilled reading only involved the direct route. Doctor and Coltheart (1980) suggested that word meanings are activated primarily by the direct route in skilled reading, with the phonological route playing a greater role in reading acquisition and the reading of exception words. Van Orden, Pennington, and Stone (1990) argue, however, that the phonological route plays a prominent role even in skilled reading. A growing body of evidence indicates that a phonological processing deficit is at the root of poor reading. For example, Perry and Ziegler (2000) noted that there is increasing consensus that poorly specified, imprecise, or inadequate phonological representations are major factors underlying reading impairment. Consistent with this notion, Shankweiler, Lundquist, Dreyer, and Dickinson (1996) found that skilled readers have a better understanding of how orthography represents the phonology of words. Thus, phonological processing ability is important for becoming a successful reader. It may be that the phonological route plays a crucial role in establishing connections between print and meaning, and that skilled readers primarily activate these direct connections during word identification.

For high print exposure participants, there was no significant difference in performance when the nonwords were orthographically legal (Experiment 1) and when the nonwords were pseudohomophones (Experiment 2). In addition, there was no pseudohomophone effect for this group of participants. This would suggest that high print exposure participants used the same word recognition processing strategies in the two experiments. One explanation would be that the high print exposure participants generally use the direct route to make word/nonword discriminations. If this were the case, performance in the two experiments would be expected to be identical, and no pseudohomophone effect should occur. For low print exposure participants, performance in Experiments 1 and 2 was quite different, and these participants did exhibit a classic pseudohomophone effect. Unlike the high print exposure participants then, these participants were affected by pseudohomophones, which suggests that they relied more on phonological information to make their decisions. That is, the low print exposure participants may generally rely on the phonological route when reading.

Another possible explanation is suggested by a recent study by Pexman, Lupker, and Jared (2001). According to these investigators, pseudohomophone foils make the lexical decision task more difficult because pseudohomophones activate the phonological representations of words. Moreover, this phonological activation feeds back to the orthographic level and activates the orthographic representations of words. To avoid positively responding to a pseudohomophone in a lexical decision task, participants have to set a more strict criterion for discriminating between words and nonwords. As a result, there is more time for feedback activation to accumulate and create competition at the orthographic level, which would produce larger word frequency and neighborhood size effects on response latencies. High print exposure participants would set a more conservative criterion than low print exposure participants because their exposure to print is greater and so their knowledge of the connection between orthography and phonology is stronger. As a result, they would not exhibit increased word frequency and neighborhood size effects.

In conclusion, facility in recognition of words in printed form is strongly tied to phonological processing, which is an important skill one must acquire to be a skilled reader. The present results suggest that differences in phonological processing skills are an important distinguishing characteristic of high print exposure and low print exposure participants.

References

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

Andrews, S. (1992). Frequency and neighborhood effects on lexical similarity or orthographic redundancy? Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 234-254.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. Psychonomic Bulletin and Review, 4, 439-461.

Bell, L. C., & Perfetti, C. A. (1994). Reading skill: Some adult comparisons. Journal of Educational Psychology, 86, 244-255.

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. Psychological Review, 102, 146-184.

Borowsky, R., Masson, M. E. J. (1999). Frequency effects and lexical access: On the interpretation of null pseudohomophone base-word frequency effects. Journal of Experimental Psychology: Human Perception and Performance, 25, 270-275.

Bräten, I., Lie, A., Andreassen, R., & Olaussen, B. S. (1999). Leisure time reading and orthographic processes in word recognition among Norwegian third-and fourth-grade students. Reading and Writing: An Interdisciplinary Journal, 11, 65-88.

Brown, J. I. (1900). Nelson-Denny Reading Test. Boston, MA: Houghton Mifflin.

Butler, B. E., Jared, D., & Hains, S. (1984). Reading skill and the use of orthographic knowledge by mature readers. Psychological Research, 46, 337-353.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhoods in visual word recognition: Cross-task comparisons. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 857-871.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. Memory & Cognition, 28, 143-153.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), Strategies of Information Processing (pp. 151-216). San Diego, CA: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), Attention and Performance VI, (pp. 535-555). Hillsdale, NJ: Erlbaum.

Davelaar, E., Coltheart, M., Besner, D., Jonasson, J. T. (1978). Phonological encoding and lexical access. Quarterly Journal of Experimental Psychology, 31, 489-507.

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. Memory & Cognition, 8, 195-209.

Elboro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. Reading and Writing, 8, 453-485.

Forster, K. I. (1976). Accessing the mental lexicon. In R. Wales & E. Walker (Eds.), New Approaches to Language Mechanisms (pp.257-287). Amsterdam: North Holland.

Forster, K. I. (1987). Form priming with masked primes: The best match hypothesis. In M. Coltheart (Ed.), Attention and Performance XII, (pp.201-219). Hillsdale, NJ: Erlbaum.

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

Forster, K.I, & Davis, C. (1991). The density constraint on form-priming in the naming task: Interference effects from a masked prime. Journal of Memory and Language, 30, 1-25.

Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory effects in lexical decision and semantic categorization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 696-713.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and False Trails. Psychological Bulletin, 123, 71-99.

Grainger, J., O'Regan, J. K., Jacobs, A. M., & Segui, J. (1989). On the role of competing word units in visual word recognition: The neighborhood frequency effect. Perception and Psychophysics, 45, 189-195.

Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. Journal of Memory and Language, 29, 228-244.

Grainger, J., & Segui, J. (1990). Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies. Perception and Psychophysics, 47, 191-198.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple-read out model. Psychological Review, 103, 518-565.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. Journal of Experimental Psychology: General, 113, 256-281.

Gordon, B. (1985). Subjective frequency and the lexical decision latency function: Implications for mechanisms of lexical access. Journal of Memory and Language, 24, 631-645.

Huntsman, L. A., & Lima, S. D. (1996). Orthographic neighborhood structure and lexical access. Journal of Psycholinguistic Research, 25, 417-429.

Jared, D., Levy, B. A., & Rayner K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from Proofreading and eye movements. Journal of

Experimental Psychology: General, 128, 219-264.

Kucera, H., & Francis, W. N. (1967). Computational Analysis of Present-Day American English. Providence, RI: Brown University Press.

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

Massoro, D. W., & Taylor, G. A. (1980). Reading ability and utilization of orthographic structure in reading. Journal of Educational Psychology, 72, 730-742.

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

McCann, R S., Besner, D., & Davelaar, E. (1988). Word recognition and identification: Do word-frequency effects reflect lexical access? Journal of Experimental Psychology: Human Perception & Performance, 14, 693-706.

McQuade, D. V. (1981). Variable reliance on phonological information in visual word recognition. Language and Speech, 24, 99-109.

Monsell, S., Patterson, K. E., Graham, A., & Hughes, C. H. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 452-467.

Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation-verification model for letter and word recognition: The word superiority effect. Psychological Review, 89, 573-594.

Paap, K. R., & Johansen, L. S. (1994). The case of the vanishing frequency effect: A retest of the verification model. Journal of Experimental Psychology: Human Perception and Performance, 20, 1129-1157.

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. Journal of Experimental Psychology: Human Perception and Performance, 24, 767-779.

Perfetti, C. A. (1985). Reading Ability. New York, NY: Oxford University Press.

Perry, C., & Ziegler, J. C. (2000). Linguistic difficulties in language and reading development constrain skilled adult reading. Memory and Cognition, 28, 739-745.

Pexman, P. M., Lupker, S. J., & Jared, D. (2001). Homophone Effects in Lexical Decision. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 139-156.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2001). Investigating the impact of feedback phonology in visual word identification. Journal of Experimental Psychology: Learning, Memory and Cognition, in press.

Pugh, K. R., Rexer, K., Katz, L. (1994). Evidence of flexible coding in visual word recognition. Journal of Experimental Psychology: Human Perception and Performance, 20, 807-825.

Sears, C. R., Hino, Y., & Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. Journal of Experimental Psychology: Human Perception and Performance, 21, 876-900.

Sears, C. R., Lupker, S. J., & Hino, Y. (1999). Orthographic neighborhood effects in perceptual identification and semantic categorization tasks: A test of the multiple read-out model. Perception & Psychophysics, 61, 1537-1554.

Sears, C. R., Hino, Y., & Lupker, S. J. (1999). Orthographic neighborhood effects in parallel distributed processing models. Canadian Journal of Experimental Psychology, 53, 220-229.

Seidenberg, M. S., Petersen, A., MacDonald, M. C., Plaut, D. C. (1996). Pseudohomophone effects and models of word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 48-62.

Siakaluk, P. D., Sears, C. R., & Lupker, S. J. (2001). Orthographic neighborhood effects in lexical decision: The effects of nonword orthographic neighborhood size. Journal of Experimental Psychology: Human Perception and Performance, in press.

Shankweiler, D., Lundquist, E., Dreyer, L. G., & Dickinson, C. C. (1996). Reading and spelling difficulties in high school students: Causes and Consequences. Reading & Writing, 8, 267-294.

Stanovich, K. E., & West R. F. (1989). Exposure to print and orthographic processing. Reading Research Quarterly, 24, 402-433.

Stanovich, K. E., Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. Memory & Cognition, 20, 51-68.

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. Journal of Experimental Psychology: Human Perception and Performance, 19, 744-774.

Van Hoy, M., & Van Orden, G. C. (2001). Pseudohomophones and word recognition. Memory & Cognition, 29, 522-529.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. Memory & Cognition, 15, 181-198.

Van Orden, G. C. (1991). Phonologic mediation is fundamental to reading. Besner, D. (Ed), Humphreys, G. W. (Ed), Basic processes in reading: Visual word recognition. (pp. 77-103).

Van Orden, G. C, Stone, G. O, Garlington, K. L., Markson, L. R., Pinnt, G., S., Simonfy, C. M., Brichetto, T. (1992). "Assembled" phonology and reading: A case study in how theoretical perspective shapes empirical investigation. Frost, R. & Katz, L. (Ed). Orthography, Phonology, Morphology, and Meaning: Advances in Psychology, Vol. 94. (pp. 249-292). Amsterdam, Netherlands: North-Holland: Lawrence Erlbaum Associates, Inc.

Van Orden, G. C., Pennington, B. F., & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. Psychological Review, 97, 488-522.

West, K. E., Stanovich, R. F., & Mitchell, H. R. (1993). Reading in the real world and its correlates. Reading Research Quarterly, 28, 35-50.

Table 1

Mean Word Frequency and Neighborhood Size for the Word Stimuli Used in Experiments 1 and 2

	Neighborhood Size	
Stimulus Characteristic	Small	Large
	Low-Frequency Words	
Word Frequency	21.6	23.6
Neighborhood Size	2.6	10.4
	High-Frequency Words	
Word Frequency	191.8	188.1
Neighborhood Size	2.9	10.6

Table 2

Mean Response Latencies (in Milliseconds) and Error Rates (in %) for the Word Stimuli in Experiment 1

	Neighborhood Size	
Word Frequency	Small	Large
	Low ART Group	
Low Frequency	568 (6.9)	548 (4.0)
High Frequency	516 (2.6)	514 (2.2)
	High ART Group	
Low Frequency	566 (5.3)	549 (3.5)
High Frequency	524 (1.4)	520 (1.8)

Note. Error rates in parenthesis ().

Table 3

Mean Response Latencies (in Milliseconds) and Error Rates (in %) for the Nonword Stimuli Used in Experiments 1 and 2

Type of Nonword	Group	
	Low ART	High ART
	Experiment 1	
Small Neighborhood	612 (3.9)	601 (3.6)
Large Neighborhood	638 (7.7)	619 (5.7)
	Experiment 2	
Unusual Pseudohomophones	579 (0.2)	566 (1.0)
Common Pseudohomophones	685 (10.0)	637 (5.5)
All Common Pseudohomophones	679 (10.1)	635 (5.8)

Note. Error rates in parenthesis ().

Table 4
Mean Response Latencies (in Milliseconds) and Error Rates (in %) for the Word Stimuli in Experiment 2

	Neighborhood Size	
Word Frequency	Small	Large
	Low ART Group	
Low Frequency	604 (10.4)	564 (6.6)
High Frequency	533 (3.0)	530 (2.6)
	High ART Group	
Low Frequency	567 (4.1)	548 (2.4)
High Frequency	529 (1.4)	524 (1.5)

Note. Error rates in parenthesis ().

Table 5

Semipartial Correlation Coefficients for the Predictor Variables Used in Experiment 3 in Analyses Using Log Kucera and Francis (1967) Word Frequency

Predictor Variable	Group	
	Low ART	High ART
	Low-frequency and High-frequency Words	
Log Word Frequency	.19[a]	.17[b]
Number of LF Neighbors	.10[c]	.07
Number of HF Neighbors	.02	.05
Bigram Frequency	.01	.01
	Low-frequency Words Only	
Log Word Frequency	.16[c]	.20[b]
Number of LF Neighbors	.15[c]	.09
Number of HF Neighbors	.01	.01
Bigram Frequency	.01	.01

Note. LF = lower frequency, HF = higher frequency. [a] = $p < .001$, [b] = $p < .01$, [c] = $p < .05$

Table 6
Semipartial Correlation Coefficients for the Predictor Variables Used in Experiment 3 in Analyses Using Subjective Frequency

	Group	
Predictor Variable	Low ART	High ART
	Low-frequency and High-frequency Words	
Subjective Frequency	.41[a]	.38[a]
Number of LF Neighbors	.10[c]	.06
Number of HF Neighbors	.03	.01
Bigram Frequency	.01	.01
	Low-frequency Words Only	
Subjective Frequency	.49[a]	.49[a]
Number of LF Neighbors	.21[a]	.16[b]
Number of HF Neighbors	.06	.01
Bigram Frequency	.01	.01

Note. LF = lower frequency, HF = higher frequency. [a] = $p < .001$, [b] = $p < .01$, [c] = $p < .05$