ELSEVIER

Neuropsychologia 43 (2005) 2068–2083

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Behavioral and neurobiological effects of printed word repetition in lexical decision and naming

Leonard Katz [a,b,*], Chang H. Lee [a,c,**], Whitney Tabor [a,b], Stephen J. Frost [a], W. Einar Mencl [a], Rebecca Sandak [a], Jay Rueckl [a,b], Kenneth R. Pugh [a,d]

[a] *Haskins Laboratories, New Haven, CT, USA*
[b] *Psychology Department, University of Connecticut, Storrs, CT 06269-1020, USA*
[c] *Department of Psychology, Pusan National University, 30 Changjeon-dong, Keumjeong-ku, Pusan 609-735, South Korea*
[d] *Yale University School of Medicine, New Haven, CT, USA*

Received 16 June 2004; received in revised form 15 March 2005; accepted 18 March 2005
Available online 22 April 2005

Abstract

A series of experiments studied the effects of repetition of printed words on (1) lexical decision (LD) and naming (NAM) behavior and (2) concomitant brain activation. It was hypothesized that subword phonological analysis (assembly) would decrease with increasing word familiarity and the greater decrease would occur in LD, a task that is believed to be less dependent on assembly than naming. As a behavioral marker of assembly, we utilized the regularity effect (the difference in response latency between words with regular versus irregular spelling-sound correspondences). In addition to repetition, stimulus familiarity was manipulated by word frequency and case alternation. Both experiments revealed an initial latency disadvantage for low frequency irregular words suggesting that assembly is the dominant process in both tasks when items are unfamiliar. As items become more familiar with repetition, the regularity effect disappeared in LD but persisted in NAM. Brain activation patterns for repeated words that were observed in fMRI paralleled the behavioral studies in showing greater reductions in activity under lexical decision than naming for regions previously identified as involved in assembly.

Keywords: Word recognition; Repetition; fMRI; Reading

The present behavioral and brain imaging experiments explore a neurocognitive explanation for how word processing becomes fluent with repetition. The experiments compare change in the recognition and naming of printed words as a function of repeated exposure, word frequency, and alternating orthographic case, that is, change as a result of differences in word familiarity.

One source of motivation for the present experiments derives from brain imaging data indicating that the brain regions prominent in word recognition are different when a relatively unfamiliar word is processed (e.g., a pseudoword or low frequency word) compared to those when a frequent word is processed. A tentative interpretation of the brain imaging data suggests that the regions prominent in processing unfamiliar words are associated with a cognitive process of assembly (e.g., grapheme–phoneme conversion and other word-internal phonological analysis), unlike the regions associated with processing familiar words (for a review of the imaging data, see Poldrack & Wagner, 2004; Pugh, Mencl, Jenner, Lee et al., 2001).

Computational models of word recognition provide a second source of motivation (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Harm & Seidenberg, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996). Paralleling the neurobiological data, these models also suggest that assembled phonology plays a diminished role in the recognition

* Corresponding author. Tel.: +1 860 486 4303; fax: +1 860 486 2760.
** Co-corresponding author.
E-mail addresses: Katz@uconn.edu (L. Katz), chleehoan@pusan.ac.kr (C.H. Lee).

0028-3932/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2005.03.022

process as a word becomes familiar (e.g., Harm, McCandliss, & Seidenberg, 2003).

The experiments we report here increase word familiarity by repeating an item within the experimental session (Scarborough, Cortese, & Scarborough, 1977). In the fMRI experiment, the effects of increased repetition are observed on the activity (specifically, blood flow) of specific brain regions that previous research suggests play a role. In the behavioral experiments, the index used to observe change in assembly is the regularity effect. The regularity effect refers to the finding that response latencies are faster to words whose spelling follows standard grapheme–phoneme conversion rules for pronunciation (e.g., *made*, *mill*) than for words with irregular or exceptional spelling-sound correspondences (e.g., *have*, *pint*) (Coltheart & Rastle, 1994).

The regularity effect is central to the behavioral experiments presented here because, when it is found, it can be taken as evidence that the word recognition process depends critically on assembled phonological information (Visser & Besner, 2001). For regular words, the phonological information produced by assembly is informative about the word's entry in the reader's phonological lexicon. However, for an irregular word, assembly can produce phonology that conflicts with the lexicon's stored pronunciation for that word, slowing processing. On the other hand, when the recognition process does not depend on assembly, regular and irregular words will not differ in their speed of lexical access because each whole printed word has a unique pronunciation (there are a very few exceptions to this, viz., the heterophonic whole words like *bow*, *read*, *wind*, etc.).

Research on printed word processing has predominantly utilized both lexical decision and naming tasks. In behavioral data, regularity effects are nearly ubiquitous in naming (although usually only for low frequency words) but are found less often in lexical decision (Hino & Lupker, 2000). This finding can be interpreted to mean that the requirement to produce overt speech in the naming task somehow promotes assembly (Frost, Katz, & Bentin, 1987; Seidenberg, Waters, Barnes, & Tannenhaus, 1984). Therefore, it was of interest to study the effects of repetition in both kinds of tasks to determine if repetition, by decreasing the use of assembly, would be more successful in attenuating the regularity effect in lexical decision than in naming. We then followed the behavioral experiments with brain imaging studies to determine if repetition attenuated activity in brain regions associated with assembly faster when the task was lexical decision than when it was naming.

As we noted, when a regularity effect occurs, it is generally for low frequency words; only rarely is it found for high frequency words (Jared, 1997), suggesting that familiarity deceases dependence on assembly. A neurobiological parallel of this attenuation can be found in brain imaging data from several laboratories, including our own. There are brain circuits whose contributions change when they process familiar, as opposed to unfamiliar, printed words. This evidence has shown that the processing of unfamiliar words is supported, in part, by a region of the temporoparietal region, more specifically, the supramarginal gyrus (Price, Winterburn, Giraud, Moore, & Noppeney, 2003; Pugh, Mencl, Jenner et al., 2000; Xu et al., 2001). The evidence indicated that this system worked in conjunction with the left hemisphere (LH) inferior frontal gyrus (IFG) and related frontal lobe areas. The processing of familiar words, in contrast, was found to be supported by a more ventral system located in the LH occipitotemporal area, including the fusiform gyrus, and more anterior sites that include middle temporal gyrus (MTG) (Sandak et al., 2004). For convenience, we will refer to the three main circuits involved in printed word identification as IFG, the temporoparietal (TP), and the ventral.

These and other behavioral and imaging data led us to hypothesize that the effect of increasing the familiarity of a word by means of repetition should be modulated by task differences (Fiebach, Friederici, Mueller, & von Cramon, 2002; Helenius, Tarkiainen, Cornelissen, Hansen, & Salmelin, 1999; Helenius, Uutela, & Hari, 1999; Pugh et al., 1997; Pugh, Mencl, Shaywitz et al., 2000; Pugh, Mencl, Jenner, Lee et al., 2001; Pugh, Mencl, Jenner, Katz et al., 2001; Rumsey et al., 1997; Shaywitz et al., 1998; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). In lexical decision, the effect of repeating a stimulus word should reduce phonological assembly, reduce IFG activation, and move the dominant locus of processing from the TP to the ventral circuit. The behavioral concomitant of reduced activation in IFG and the TP circuit should be a reduced regularity effect (Poldrack & Wagner, 2004). In contrast, when the task is naming, repetition may not influence the regularity effect as strongly because the articulatory demands of the task require IFG activation to be maintained, which may also sustain activation of the TP circuit. Because we hypothesize that these circuits are important for assembly, their maintenance under repetition should mitigate reduction of the regularity effect.

Thus, the present research addresses the following specific questions about the effects of repeated experience on printed word processing. First, does repetition induce a reduction in assembly, as indexed behaviorally by the regularity effect? Secondly, is this reduction greater for the lexical decision task (which does not require articulation) than for the naming task (which does)? Third, will repetition also produce a reduction in activity in IFG and the TP circuit (regions associated with assembly) with a concomitant increase in ventral circuit efficiency (as the neurobiological model proposes) and will this reduction be greater for the lexical decision task than naming?

1. Experiment 1

Early in learning, when a word is relatively unfamiliar, the dominant process of word recognition will be subword phonological analysis, i.e., assembly. Assembly may slow the recognition of irregular words relative to regular words. However, as a word is repeated, the dominance of assembled phonological analysis will diminish and the regularity effect that resulted from it will be reduced. This was the interpretation of Visser and Besner (2001) who found just such a reduction in the regularity effect when the words were repeated within a naming task.

In the present experiment, participants received four repeated blocks of the same list of words and nonwords. Half of the words had irregular and half regular spelling-to-sound correspondences. In addition, we introduced two additional manipulations of familiarity. Mixed case was used to manipulate a word's orthographic familiarity. One group of participants saw the stimulus items in mixed case (e.g., *mInT*), an orthographic form that was unfamiliar. With a mixed case word, a reader is confronted with a stimulus for which he/she has no memory-based orthographic representation (except at a more abstract level) and, therefore, the reader should be more likely to engage in assembly. Thus, we should see a larger regularity effect for mixed case words. Further, it might persist longer over repetitions and might appear even for high frequency mixed case words. However, with enough repetition, processing based on assembled phonology should give way to processing whose lexical access units are of a larger orthographic grain-size (Ziegler & Goswami, 2005). Then, they should be recognized as quickly as items that were initially more familiar.

An additional manipulation of familiarity was word frequency. Familiarity, defined as word frequency, should be somewhat similar to the effect of repetition; words of high frequency have simply been experienced more often (although with a different distribution and context for repetitions). Because of the participant's past reading experience, he/she should be more likely to depend on assembly for infrequently experienced words than for high frequency words.

1.1. Method

1.1.1. Subjects

Seventy-six undergraduate students at the University of Connecticut participated in the experiment for course credits. All participants reported normal or corrected-to-normal vision. Each subject was randomly assigned to one of two counterbalanced groups, resulting in 38 subjects per group.

1.1.2. Materials

Eighty-eight words and the same number of nonwords were selected to construct the stimulus list. Half of the words were high frequency (333.10 tokens per million, from the CELEX database, Baayen, Piepenbrock, & van Rijn, 1993), and the other half were low frequency words (9.99). Half the words in each frequency group were regular words and half irregular words. The length and average frequency of regular and irregular words were matched. The frequency of these four subgroups were: 329.2, for high frequency regular words; 336.9 for high frequency irregular words; 11.0 for low frequency regular words; 8.5 for low frequency irregular words. Two stimulus lists, with the same words and nonwords in each list, were created in order to investigate the effect of mixing case. Half the stimuli in each list were printed in mixed case and half in uniform case. Mixed case words began with a lowercase letter and alternated the case between successive letters (e.g., bEaCh). Uniform case words were printed all in lowercase letters. The two lists differed in that words that were mixed case in one list were composed in uniform case in the other and vice versa. Nonwords were orthographically legal and pronounceable and were created by changing one letter of a real word that was not one of the word stimuli. Half of the nonwords were also printed in mixed case and half in uniform case. An additional 16 words and nonwords were created as the practice stimuli. The stimuli for Experiment 1 are listed in Appendix A.

1.1.3. Procedure

The presentation of stimuli and data recording were controlled by the experiment software, E-prime (Schneider, Eschman, & Zuccolotto, 2002). Participants sat in front of a computer monitor at a viewing distance of about 60 cm. Stimuli were presented in the center of the monitor screen and remained on until the subject responded. The stimuli were presented as white characters on the dark background of the screen. The intertrial interval was 1 s. Participants were told that the stimuli they saw in the first block would be repeated three more times for a total of four blocks. Participants were required to decide the lexicality of each letter string as quickly and as accurately as possible. Subjects pressed one of two telegraph keys to indicate a word or nonword decision. Subjects were given a short break of about 30 s after each block. An experimenter sat beside the subject throughout the experiment.

1.2. Results

1.2.1. Analysis of words: RT

Response latencies less than 100 ms were discarded. These outliers were less than 0.7% of all responses. For the reaction time (RT) analysis, trials on which an error occurred were also discarded. There was no significant main effect or interaction involving the counterbalanced lists (in fact, all Fs were small), enabling us to combine the data for the two lists. Analyses of variance (ANOVAs) were conducted once with subjects as the unit of analysis and once with items (stimuli).

For the subjects analysis, the design was a $2 \times 2 \times 2 \times 4$ (frequency × case × regularity × block) repeated-measures design. All main effects were statistically significant: block, $F(3, 228) = 13.08$, MSE = 15986.70, $p < .001$; frequency, $F(1, 76) = 390.77$, MSE = 5061.17, $p < .0001$; case, $F(1, 76) = 115.48$, MSE = 5646.04, $p < .0001$; regularity, $F(1, 76) = 15.12$, MSE = 2526.11, $p < .001$. Each of the two-way interactions between block and the other variables were all statistically significant: block × case, $F(3, 228) = 11.76$, MSE = 2967.35, $p < .001$; block × frequency, $F(3,$

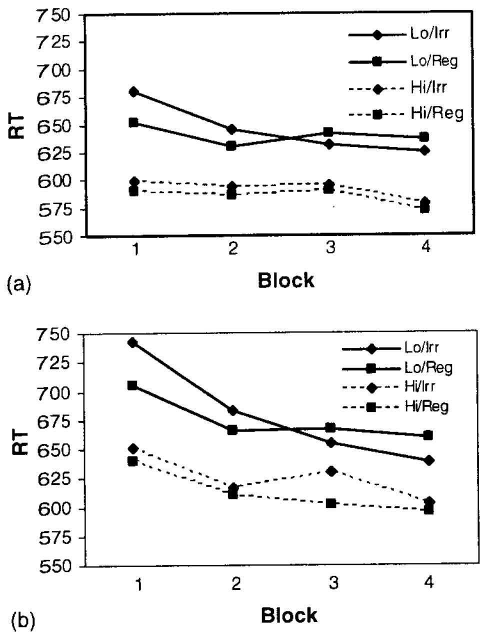

Fig. 1. Mean lexical decision latencies for frequency and regularity in (a) uniform and (b) mixed case.

228) = 9.96, MSE = 2665.46, $p < .001$; block × regularity, $F(3, 228) = 6.25$, MSE = 2604.75, $p < .001$. Importantly, the three-way interaction, block × frequency × regularity, was statistically significant, $F(3, 228) = 8.98$, $p < .001$, MSE = 2612.27. Inspection of the data in Fig. 1 suggests that the regularity effect for low frequency words in the early blocks disappeared by the third block.

A stimulus analysis was also performed. It was designed to assess and adjust for a potentially damaging stimulus confound, orthographic neighborhood density. Orthographic neighborhood density of the low frequency irregular condition was smaller than the other three word conditions, a fact that could, in principle, account for the initially slower responses in that condition; more neighbors are usually facilitative (Andrews, 1997). Mean neighborhood size (and S.D.), as measured by Coltheart's *N* (Coltheart, Davelaar, Jonasson, & Besner, 1977), were: nonwords, 7.88 (4.7), high frequency regular, 9.59 (5.2); low frequency regular, 8.18 (4.4), high frequency irregular, 8.95 (4.4); low frequency irregular, 6.27 (4.4).

The stimulus analysis in which neighborhood density was entered as a covariate failed to find any significant effects of *N*. On the contrary, mean squares for terms with *N* were small. There were significant effects for case, $F(1, 83) = 6.00$, MSE = 8975.30, $p < .02$, and block × case, $F(3, 249) = 3.12$, MSE = 1070.27, $p < .03$. Between-items, there was a significant effect for frequency, $F(1, 83) = 46.43$, MSE = 13693.55, $p < .0001$, but no significant main effect of regularity, unlike the subjects analysis. There were also significant effects for block, $F(3, 249) = 23.11$, $p < .0001$; block × frequency, $F(3, 249) = 7.21$, $p < .001$; block × regularity, $F(3, 249) = 4.50$, $p < .004$, and importantly, block × frequency × regularity, $F(3, 249) = 5.58$, $p < .001$, all on the same error term, MSE = 1196.25.

1.2.2. Analysis of words: errors

Analyses of errors paralleled analyses for response latencies. For words, the average error rate (which included outliers) was 6%. Low frequency error rate decreased over blocks from 12 to 9%; for high frequency words, the range was fairly constant at 3–4%. Error rates by condition were similar to the pattern for response latencies, although for errors, regularity and only one of the two-way interactions reached statistical significance. A subjects ANOVA showed statistically significant main effects: for frequency, $F(1, 76) = 255.80$, MSE = .019, $p < .0001$, for case, $F(1, 76) = 17.28$, MSE = .010, $p < .0001$, and for block, $F(3, 228) = 3.08$, MSE = .004, $p < .03$. The frequency × regularity interaction was significant, $F(1, 76) = 5.35$, $p < .03$, and importantly, the three-way interaction of block × frequency × regularity was also significant, $F(3, 228) = 7.96$, MSE = .03709, $p < .001$. It paralleled the three-way interaction for RT in that differences that between regular and irregular low frequency words became negligible by block 3.

1.2.3. Analysis of nonwords: RT

Half the nonwords had been presented to participants in uniform case and half in mixed case. Mean RT for nonwords in block 1 was 742 ms for uniform case and 786 ms for mixed case. Mean RT decreased consistently over blocks and was 654 ms for uniform case and 669 ms for mixed case; the difference between mixed and uniform case nonwords decreasing over blocks. The main effects of both case and block were statistically significant in a subjects ANOVA: for case, $F(1, 76) = 46.68$, MSE = 1864.06, $p < .001$, and for block, $F(3, 228) = 60.82$, MSE = 4993.77, $p < .001$. The two-way interaction was also significant, $F(3, 228) = 9.11$, MSE = 805.31, $p < .001$. An items analysis gave substantially the same results: there were significant effects of block, $F(3, 258) = 265.37$, MSE = 1308.82, $p < .0001$; case, $F(1, 86) = 10.66$, $p < .002$; and block × case, $F(3, 258) = 7.00$, MSE = 1191.15, $p < .01$.

1.2.4. Analysis of nonwords: errors

For nonwords, only the main effect of blocks was significant in the subjects ANOVA: $F(3, 228) = 4.79$, MSE = .002, $p < .003$. Unlike the nearly linear decrease in errors over trials that was observed for words, errors for nonwords were fairly constant over blocks. The items ANOVA gave the same results: only block was significant, $F(3, 261) = 6.53$, MSE = .001, $p < .01$.

1.3. Discussion

Repetition generally decreased recognition RT, with the larger decreases between the first and second presentations for mixed case words and low frequency words (Fig. 1). With regard to a regularity effect, none was observed for high frequency words, supporting the interpretation that these words are not recognized by means of assembly. Low frequency irregular words did show an initial processing inferiority relative to low frequency regular words but after two repetitions they were recognized no more slowly. This convergence of irregular and regular word recognition times with increasing experience demonstrates for lexical decision what Visser and Besner (2001) showed for naming, viz., subword spelling-to-sound inconsistencies become less effective with increasing familiarity, reflecting, presumably, the reduced dominance of word-analytic processing.

Overall, mixed case stimuli were recognized more slowly than uniform case stimuli. This general mixed case disadvantage was attenuated somewhat by repetition. Nevertheless, the regularity effect for low frequency words was not markedly stronger for mixed case than that for uniform case, contrary to expectations. In addition, there was no regularity effect for high frequency mixed case words. Herdman, Chernecki, and Norris (1999) did find a regularity effect on high frequency mixed case words in the naming paradigm. However, as we have suggested above, the use of phonological assembly may be stronger in naming than lexical decision. We explored naming in the second experiment.

First, however, we discuss some artifactual explanations of the results. It might be argued that the regularity effect decreased over time because subjects simply exchanged the irregular pronunciation of an ambivalent rime for the regular one (e.g., the rime—*int* was effectively read as only it is pronounced in *pint*) and, because it was now the dominant (or only) response, irregular words like *pint* no longer produced any conflict in phonological lexical access. If this hypothesis is accurate, then we should also see the same reduction of the regularity effect in the next experiment, naming. (To anticipate the result of Experiment 2: we did not see such a reduction.) We did run a related lexical decision experiment (not reported here in full because of space limitations), which was designed like Experiment 1 but included a regular word counterpart for every irregular word. That is, for every *pint* type of word there was a *mint* type. Therefore, the subject was exposed equally to both pronunciations of the ambivalent rime and, therefore, neither would have an advantage. The results of this study were essentially the same as in Experiment 1; a significant initial regularity effect became nil quickly with repetition. Thus, the reduction of the regularity effect cannot be attributed to substituting the irregular pronunciation for the regular.

A second alternative explanation concerns the nature of learning in lexical decision. It is possible that subjects were merely learning a discrimination task in which a simple response (yes or no) was learned to each stimulus as it was repeated. The reduction in the regularity effect would then be due to subjects reading the stimuli not as words and nonwords but as cues for a response. Such a mechanism is less plausible (because learning a new response to each item seems more effortful than relying on one's pre-experimental lexical knowledge for the appropriate response) but present data cannot rule it out.

2. Experiment 2

Experiment 2 used the word stimuli of the lexical decision experiment in a naming paradigm. We conjectured that the rapid elimination of the regularity effect in lexical decision (interpreted as a decreased dependence on assembly) would not occur in naming, a task in which phonological analysis of the printed word is thought to be more dominant.

Perhaps, it is not surprising that the naming paradigm has been found to be more sensitive than lexical decision to regularity and other phonological factors. This sensitivity may be a neurobiological consequence of the requirement to articulate. Consider the possibility that articulation maintains activation of inferior frontal gyrus and the temporoparietal circuit. These circuits are also believed to support assembly (Poldrack & Wagner, 2004; Price et al., 2003; Sandak et al., 2004; Shaywitz et al., 2002; Xu et al., 2001). When articulation is required, assembly may, in fact, be efficacious: it may be easier to articulate a target word rapidly if some of the phonology needed for the articulation process has already been activated at an earlier stage. By generating phonemic and syllabic information before the word has been uniquely identified (a dual-route theorist might say, "before lexical access"), the reader may achieve a speed advantage in naming because some of the information needed for the articulatory plan becomes available even before the word itself is fully specified. Thus, in naming, the reader may have a "head start" on articulation if the reader has already generated some of the required phonology. In contrast, in lexical decision, there is no requirement to articulate and the faster process may be the one that does not extract subword phonological information.

Thus, if a printed word is initially unfamiliar and, therefore, processed by the TP circuit, processing by IFG and TP will persist under repetition longer if the task is naming than if it is lexical decision. We should see this manifested behaviorally as a regularity effect that tends to persist in spite of repeated presentations. However, even in naming, the regularity effect should eventually disappear, as a word continues to become more and more familiar. Visser and Besner (2001) found an initial regularity effect in a naming task but the effect was reduced after a single repetition. Thus, the question is: will repetition have the effect of eliminating the regularity effect in naming as rapidly as it did in lexical decision? If not, an explanation of the difference may reside in the efficacy of assembly for naming and the greater neurological/phonological activation of the dorsal circuit (i.e., IFG and TP) that the naming task requires.

As in Experiment 1, we decreased the orthographic familiarity of half the stimuli by printing them in mixed case. When Herdman et al. (1999) printed words in mixed case, they found a regularity effect for naming even for high frequency words. As usual, they found no regularity effect on these words when printed in uniform case. We also expected a stronger initial regularity effect for the words made less orthographically familiar by case mixing but there should be an eventual reduction in the effect with repetition. Familiarity was also manipulated by means of word frequency; it was expected that a regularity effect (an index of assembly) would be more prominent for low frequency words.

2.1. Method

Experiment 2 parallels Experiment 1 in design. It has the same 88 words as those in Experiment 1 and, as in Experiment 1, they are repeated in four consecutive blocks. No nonwords were presented.

2.1.1. Subjects

Forty-four undergraduate students at the University of Connecticut participated in the experiment for course credits. All participants normal or corrected-to-normal vision. Each subject was quasi-randomly assigned to one of two groups counterbalanced for mixed case, resulting in 22 subjects per group.

2.1.2. Procedure

The procedure was similar to Experiment 1 except that participants were instructed to speak the word as quickly as possible. An experimenter sat beside the participant throughout the experiment and noted errors.

2.2. Results

Response latencies less than 100 ms were discarded as outliers. These outliers composed of less than 0.5% of all responses. An additional 2.2% of responses (stammers, incorrect pronunciations, etc.) were classified as errors.

2.2.1. Analysis of RT

For the reaction time analysis, trials on which an error occurred were discarded. Fig. 2a presents RT for uniform case and Fig. 2b, for mixed case.

Inspection of Fig. 2a and b suggests that naming of mixed case words was initially slower than uniform case but, by block 4, was just as fast. Low frequency words, although slower than high frequency words, decreased faster from block 1 to block 4. Importantly, the figures suggest, as predicted, that the regularity effect did not disappear after two blocks (as occurred in lexical decision) but persists over four trials.

An ANOVA based on subjects as units of analysis examined list × frequency × case × regularity × block with all but list as repeated-measures factors. There were two marginally significant four-way interactions involving counterbalanced lists. A third four-way interaction with list was more strongly significant (frequency × case × regularity × list), $F(1,42) = 31.57, p < .0001$, MSE = 2953.79. Inspection of the means for this interaction indicated that RTs were particularly long for low frequency irregular uniform case words in list 2. However, with that exception, the pattern of responding in that block was otherwise consistent with list 1. For this reason and because no effects involving list were significant

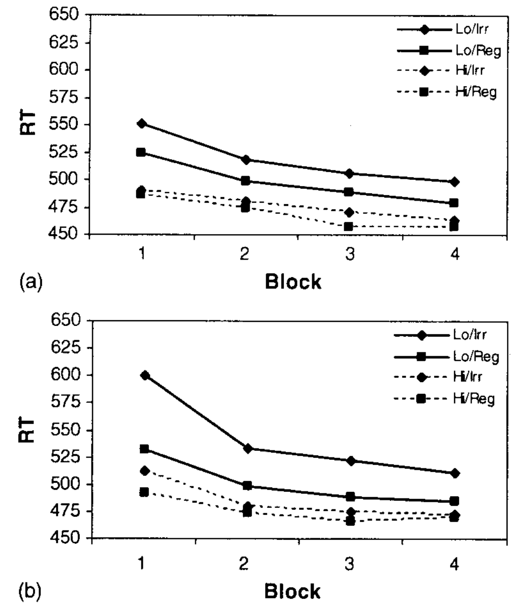

Fig. 2. Mean naming latencies for frequency and regularity in (a) uniform and (b) mixed case.

in the stimulus ANOVA (see below), we suggest that we can ignore list differences without substantive distortion.

All main effects were statistically significant. For frequency, $F(1, 42) = 244.56$, $p < .0001$, MSE = 2133.94; for case, $F(1, 42) = 30.93$, $p < .0001$, MSE = 1188.64; for regularity, $F(1, 42) = 53.04$, $p < .001$, MSE = 2509.54; for block, $F(3, 126) = 35.28$, $p < .001$, MSE = 3846.05. Also significant were the two-way interactions between case × regularity, $F(1, 42) = 18.47$, MSE = 574.09, $p < .0001$ and frequency × regularity, $F(1, 42) = 33.75$, MSE = 1299.60, $p < .0001$. Fig. 2 shows larger regularity effects for mixed case words and low frequency words.

With regard to change in RT over blocks, mixed case words had a faster initial decrease with repetition, as did low frequency words. Additionally, the regularity effect was stronger in the initial blocks. For the interaction block × case, $F(3, 126) = 8.60$, MSE = 599.17, $p < .001$; for the interaction block × frequency, $F(3, 126) = 20.31$, MSE = 669.08, $p < .001$; and for the interaction block × regularity, $F(3, 126) = 5.76$, MSE = 735.48, $p < .001$. Finally, the three-way interaction block × case × regularity was significant, $F(3, 42) = 5.50$, MSE = 587.70, $p < .001$; Fig. 2 suggests that the regularity effect was more constant over blocks for uniform case words than for mixed case words. For the latter, there was a large initial regularity effect that decreased by the second block (although it still remained larger than the regularity effect for uniform case, throughout).

Although the interaction for block × case × frequency × regularity was not significant, there was a priori interest in testing for an initial regularity effect on high frequency words. Herdman et al. (1999) had found such an effect on high frequency words, but only when they had been printed in mixed case; no regularity effect was found for high frequency uniform case. For the present mixed case data, we found a 20 ms difference between regular and irregular high frequency words in block 1 that was significant, $t(43)=2.1$, $p<.05$. The difference for uniform case words was small, 4 ms, and nonsignificant. In contrast to high frequency words, the regularity effects in block 1 for low frequency words were substantial, as can be seen in Fig. 2. For mixed case, the difference was 68 ms, $t(43)=5.49$, $p<.0001$, and for uniform case, although smaller at 27 ms, it was still significant, $t(43)=2.69$, $p<.02$. The difference between the two is also significant, $t(43)=3.71$, $p<.001$. By block 4, there was no significant regularity effect for high frequency words. However, for low frequency words, significant regularity effects still held for both mixed case $t(43)=2.74$, $p<.01$ and for uniform case $t(43)=2.65$, $p<.02$.

A stimulus analysis of covariance, with RT adjusted for neighborhood density, *N*, produced a marginally significant effect of *N*, $F(1, 83)=3.90$, MSE = 8738.87, $p=.051$. In common with the subjects analysis, there were significant effects of: frequency, $F(1, 83)=24.20$, $p<.0001$; regularity, $F(1, 83)=5.99$, $p<.02$, both with MSE = 8738.87; block, $F(3, 249)=26.24$, $p<.0001$; block × frequency, $F(3, 249)=7.89$, $p<.001$; block × case, $F(3, 249)=2.89$, $p<.04$. In contrast to the subjects analysis, there were no significant stimulus analysis effects for frequency × regularity, case × regularity, and block × case × regularity. Additionally, block × regularity, which was significant in the subjects analysis was marginally nonsignificant in the stimulus analysis: $F(3, 249)=2.21$, $p=.09$. We interpret this as somewhat weak evidence suggesting some attenuation of the regularity effect over blocks.

2.2.2. *Analysis of errors*

Error rates were low, averaging 1.5% overall and ranging from averages of 4 to 0% for the 32 combinations of block, case, frequency, and regularity. Analyses of variance on errors based on subjects and stimuli produced only one significant effect in common. For block × frequency, $F(3, 126)=3.31$, MSE = .001, $p<.03$, for subjects and $F(3, 252)=4.45$, MSE = .0009, $p<.01$. Errors for low frequency words decreased faster over blocks than for high frequency words.

2.3. *Discussion*

In contrast to lexical decision, the regularity effect in naming persisted so that even on the fourth exposure, significant differences between regular and irregular words remained. Because the regularity effect is a mark of assembled phonological processing, it appears that subjects continued to use assembly in the naming task after they had adopted, in lexical decision, an addressed routine (i.e., in which the recognition process is based on the whole word's orthographic patterns). Case mixing further promoted the use of assembly; the regularity effect was much larger for mixed case words than for uniform case. Indeed, in naming, there was a suggestion that assembled processing occurred even on high frequency mixed case words, although only on a word's first encounter (a result similar to Herdman et al., 1999).

It would be expected that case mixing would amplify the use of assembly in lexical decision also. However, that there was no such significant interaction between case mixing and regularity in Experiment 1 (only main effects). Perhaps the absence of a significant case × regularity term, there only reflects a lack of power (the differences were in the direction of a greater regularity effect for mixed case words) but it should be noted that the absence of a significant increased regularity effect for mixed case words in lexical decision is not consistent with our assumption of the regularity effect as a marker of assembled processing. Nevertheless, the bulk of the evidence from Experiments 1 and 2 suggests that, in lexical decision, the word recognition process changes quickly under repetition from assembly to a process that does not require pervasive grapheme–phoneme translation whereas, for naming, assembly persists longer.

Is it possible that the difference between tasks in the persistence of the regularity effect is an artifact, due to our design difference, which used nonwords in the lexical decision task but not in the naming task? We suggest that the absence of nonwords in the naming experiment makes for a more conservative comparison between tasks for a simple reason: the literature shows that the absence of nonwords *decreases*, not increases, phonological effects for words (Carello, Turvey, & Lukatela, 1992; Monsell et al., 1992; Tabossi & Laghi, 1992; Van-Orden, Pennington, & Stone, 1990). If nonwords had been included in naming, the effect of regularity would, presumably, have been even stronger. In addition, we have recently completed studies in which we have included nonwords (both in standard naming and in go-nogo naming) and we find essentially the same result: increased persistence in the regularity effect in naming compared to lexical decision.

We have speculated elsewhere on the stability of the regularity effect in lexical decision (Pugh et al., 1997; Pugh, Mencl, Jenner, Lee et al., 2001; Pugh, Mencl, Jenner, Katz et al., 2001). Whether one sees a regularity effect or not in group-averaged data may be a function of how different individuals contribute to the average. That stronger regularity effects have been seen for younger readers (Seidenberg, 1992) and less skilled readers (Bruck, 1990) suggests that there may be individual differences even within the population our present sample was drawn from. In a direct assessment of a college student sample, Pugh et al. (1997) showed that the size of a subject's regularity effect in lexical decision was correlated with the extent of the subject's bi-hemispheric activation in IFG, a region that is involved in speech production (among other functions).

Thus, individual variation may have more impact in the lexical decision task, where phonological assembly may be optional for a subject, thereby allowing a variety of processing styles to be expressed. In contrast, in naming, task demands may induce nearly all readers to engage in assembly. We have discussed above the possible advantage of assembly in naming because it can generate phonology prior to full lexical activation, thereby priming articulation. In lexical decision, with no need for articulation, phonology generated in assembly may offer no advantage. Because models of word recognition allow for trade-offs among various information pathways involved in the word recognition process (trade-offs that depend on characteristics of the words and characteristics of the reader), Seidenberg (1992) has characterized the issue as a "division of labor" problem. In addition to characteristics of the stimuli and the reader, Experiments 1 and 2 suggest that task demands (i.e., the kind of reading required) is yet another factor that enters into the calculations for the division of labor. We will offer neurobiological evidence for this claim in a final experiment.

3. Experiment 3

In the two tasks studied in Experiments 1 and 2, there were both common and distinct effects of repetition. For both lexical decision and naming, latencies decreased over repetitions. However, phonological assembly (as indexed by the regularity effect) persisted over repetitions in naming but not lexical decision. In Experiment 3, we tried to identify both common and distinct neurobiological signatures of repetition in the two tasks.

Activation of inferior frontal gyrus is an obligatory consequence of articulating the printed word. In addition, activation of IFG and related temporoparietal sites have been associated with phonological assembly (Fiez, Balota, Raichle, & Petersen, 1999; Herbster, Mintun, Nebes, & Becker, 1997; Poldrack & Wagner, 2004; Price et al., 2003; Sandak et al., 2004; Shaywitz et al., 2002; Xu et al., 2001). It may be the case that IFG activation, per se, encourages assembly or, as we suggested, participants in the naming task can articulate faster if they strategically utilize assembly to provide early phonological information. Either way, IFG activation should be strong in the naming task and should be relatively slow to decrease with repetition. In lexical decision, with no requirement to articulate, IFG activation should be reduced. Also, repetition in both tasks should have the effect of reducing activity in IFG and temporoparietal sites further as the recognition process grows less dependent on phonological assembly.

In order to test the hypothesis that the different behavioral effects of repetition on lexical decision and naming are associated with different responses in the temporoparietal and anterior word recognition circuits, we conducted an fMRI experiment. Given the many design constraints we were faced with, efficiency in design was a major consideration. Because of this, we did not include the factor of regularity-consistency in our fMRI design. It was not necessary to vary regularity in the imaging experiment because the neurobiological effects of the two variables of interest, viz., repetition and task, could be observed in the activity of the targeted brain regions of interest. Thus, no overt behavioral index was necessary. Indeed, in evaluating the need for a regularity manipulation, we had serious concerns about weakening the power of the imaging study by subdividing the limited time available for testing a subject into too few trials per condition. To create a comparable context for the imaging experiment, however, the stimuli employed in Experiment 3 were similar in frequency, length, and grammatical class to those employed in Experiments 1 and 2 and included both irregular as well as regular words.

Before obtaining fMRI images, we trained each participant, away from the magnet, on sets of items which were repeated in three successive blocks. Both lexical decision and naming training were given to each participant, each task immediately prior to scanning, with different stimuli in the two tasks. Then fMRI scans were obtained for those repeated items as well as for new items that the subject had not seen in training. During scanning, we contrasted blocks of repeated items with blocks of new items. In addition, there was a control task that required responding to nonverbal character displays.

Recall that in Experiments 1 and 2, the regularity effect disappeared by the third repetition trial in lexical decision but not in naming. Therefore, we expected that thrice repeated tokens in Experiment 3 would be sufficient to map out the underlying neural changes that mediated the performance differences observed in Experiments 1 and 2. We expected to find that IFG, and related dorsal sites (e.g., supramarginal gyrus, SMG) would remain relatively active even for repeated items in the naming task. Sites that were also expected to remain active in naming but not lexical decision included the supplementary motor area (SMA), and the cerebellum; both systems are known to be involved in the control of speech. In contrast, we expected to find that activity in these areas would diminish more substantially with repetition in lexical decision, as reliance on assembly was reduced or eliminated. Moreover, other brain regions associated with speech articulation should also show reductions in activity.

In contrast to the operation of the dorsal circuit, a second circuit (which we call the ventral circuit) has been found to support word recognition for familiar words. It consists mainly of occipitotemporal and middle/inferior temporal sites, (Cohen et al., 2002; Shaywitz et al., 2002; Tarkiainen et al., 1999). With regard to the ventral circuit, particularly at sites in medial temporal gyrus and inferior temporal gyrus (which are thought to be associated with lexical-semantic processing, Sandak et al., 2004), we expected activation either to increase with repetition or – what would be functionally equivalent – to remain constant while TP processing decreased so that ventral processing grew relatively stronger with repetition. In effect, ventral circuit processing should become more prominent with repetition. Although there is evidence to identify the dorsal circuit's cognitive function with phonological assembly, it is premature to identify the ventral circuit as the addressed routine for lexical access.

3.1. Method

3.1.1. Subjects

Eighteen neurologically normal participants (12 males and 6 females) with no history of neurological impairment performed both naming and lexical decision tasks in exchange for payment.

3.1.2. Materials

Twelve words were chosen for repetition in lexical decision and 24 in naming. They were monosyllabic and low to mid-frequency. Even though regularity was not formally manipulated, the stimulus set contained both regular/consistent and irregular/inconsistent words. Nonwords used as foils in the lexical decision task were also monosyllabic and were derived from word stimuli (as in Experiments 1 and 2). Specifically, mean (and S.D.) Kucera–Francis frequencies, per million, and neighborhood N (Coltheart et al., 1977) per condition, respectively, were as follows: lexical decision, repeated words, 12.91 (19.0), 8.33 (3.7); unrepeated words, 7.29 (8.8), 5.25 (3.8). For repeated nonwords, mean N (and S.D.) were 6.08, (4.4); unrepeated nonwords, 6.93 (4.6). For naming, frequency (and S.D.) for repeated words were 9.66 (16.6), 8.25 (3.7) and for unrepeated words, 9.59 (9.5) and 6.5 (4.2). Although the purpose of Experiment 3 was to study the neurobiological effects of repetition on lexical decision and naming and was not designed to study the regularity effect, the stimuli of Experiment 3 contained some irregular as well as regular words; 6 out of the 12 repeated words in lexical decision were irregular (according to Connecticut pronunciation) as were 7 out of the 24 repeated words in naming.

3.1.3. Procedure

Stimulus presentation was controlled using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). For lexical decision, response latencies were recorded using a fiberoptic device. During the training phase (prior to scanning) tokens were presented for 1 s duration with a stimulus-onset asynchrony (SOA) of 3 s. During scanning, stimuli were presented at two quasi-random SOAs, 4 and 5 s; we employed this jittering technique to allow for a possible event-related analysis (Miezin, Maccotta, Ollinger, Petersen, & Buckner, 2000). (However, the design was not optimized for event-related analysis and preliminary examination of the data determined that a block analysis was more reliable.) Five runs of lexical decision and four runs of naming were performed; the difference in number of runs was designed to equate the two tasks for scanning time. Total duration was approximately 45 min. Within each run we alternated between baseline blocks lasting 18 s and experimental blocks lasting 54 s. Within a given experimental block, items were either exclusively new or repeated. Each block had 12 trials consisting of 12 words in naming and 6 words and 6 nonwords in lexical decision, quasi-randomly presented.

Each stimulus from the list of repeated stimuli was seen once in each run. In lexical decision, the subject was told to respond as quickly as possible while maintaining accuracy by pressing one button for words and the other button for nonwords. Responses were made with the right hand only, using the middle and index fingers. For the lexical decision's baseline task, the subject responded to strings of slash marks, some with asterisks embedded (e.g., //*/) by pressing a key when the asterisk was embedded. In naming, subjects were instructed to pronounce the word presented as quickly as possible while maintaining accuracy. An overt naming response was required in the naming task. In pilot work, we have found stronger activity effects from overt articulation compared to silent naming. Problems due to movement in overt naming are small (see below). For the baseline task in naming, the subject responded to all strings of slashes and asterisks by speaking the same response, i.e., *pop*. Different baseline tasks for LD and naming were employed to factor out gross motor factors and decisional factors that would complicate the detection of both task invariant and task specific lexical processing effects. The ordering of lexical decision and naming was counterbalanced, half the subjects receiving the order LD NAM LD NAM, etc., and the other half, NAM LD NAM LD, etc.

3.1.4. Image acquisition

Functional imaging was performed on a GE 1.5 T Signa MR imaging system with an echo-planar imaging (EPI) gradient echo sequence (flip angle, 80°; echo time, 50 ms; repetition time, 2000 ms). One hundred and fifty-four images (20 axial-oblique slices, 6 mm thick, no gap) per slice location were collected for a total of 1386 images across the 9 runs. Across the five lexical decision runs, a total of 270 images were collected for the repeated and novel conditions and 225 images were collected for the baseline condition. Across the four naming runs, a total of 216 images were collected for the repeated and novel conditions and 180 images were collected for the baseline condition.

Because of the possibility of movement artifacts associated with overt speech production in the scanner, we removed movement-related activation as part of the motion-correction procedure. In addition, several empirical tests, both in our laboratory and elsewhere, (Sandak et al., 2004; Huang, Carr, & Cao, 2001; Kircher, Brammer, Levelt, Bartels, & McGuire, 2004; Palmer et al., 2001) indicate that, with careful subject instruction, head immobilization, and (minimal) practice, data are relatively free from artifacts and directly comparable to images acquired during silent periods.

3.1.5. Preprocessing

Data analysis was performed using software written in MATLAB (MathWorks, Natick, MA). Prior to analysis, the images from each run were motion corrected for three translation directions and for the three possible rotations using the SPM-96 program (Friston et al., 1995). The images were spatially filtered using a Gaussian filter of Full Width at Half Maximum (FWHM) of 3.125 mm. In the temporal domain, a high pass filter was applied to remove the drift at frequency lower than twice the period of activation paradigm (Skudlarski, Constable, & Gore, 1999).

Activated pixels were detected by comparing the images for each task to the baseline task using a split Student's t-test. The t-test was calculated for each series of images separately and later averaged. These t-values were used as a derived measure of the signal change at each voxel, relative to its own intrinsic noise variability. These subject level maps provide the base dependent measures from which the Composite maps and intersect maps (see below) were then created. Lastly, anatomic images and single-subject activation maps were transformed into a proportional three-dimensional grid (Talairach & Tournoux, 1988). This was performed first by in-plane transformation and then by slice interpolation into the 10 most superior slices of Talairach space, centered at $z = +69$, +60, +51, +42, +33, +23, +14, +5, −5, and −16, respectively.

3.1.6. Image analysis

Statistical activation maps were made using the standard univariate analysis method of applying the General Linear Model on a voxel by voxel basis. Effect maps were generated by voxel-wise comparisons – simple subtractions and other specific linear combinations – of the subject activation maps, and significance levels were assessed both by permutation and normal distribution tests. Permutation closely mirrored the results from normal distribution tests and, therefore, all data presented in this paper are the results from the normal distribution tests described below. For each complex effect of interest, a standard linear contrast (Tabachnick & Fidell, 2001) was computed across subjects. This procedure generated a single value for each voxel that was determined by the weighted comparison of one set of one or more tasks against another set. Importantly, under the null hypothesis of no effect, the expected value of this contrast was equal to zero. The extent to which the contrast value reliably deviated from zero was then assessed by normal distribution significance tests. At each voxel, an F-statistic was calculated across subjects, comparing the observed value of C to zero. Using the appropriate error term estimated from the within-condition variability, a p-value was generated for the significance of this effect. This implemented the repeated-measures analysis of variance (ANOVA) with planned comparisons. Finally, maps displaying logical combinations of effects (intersection analyses) were created by identifying voxels that follow a specific pattern of effects across a set of individual contrasts.

3.2. Results and discussion

Average response accuracy in lexical decision for repeated items was above 98% and for new words and nonwords, accuracy was 93 and 97%, respectively. In lexical decision, reaction times for repeated words, repeated nonwords, new words, and new nonwords had means (and S.D.), respectively, of 688 (14.2), 739 (19.9), 767 (17.4), and 795 (22.8). For naming, although neither accuracy nor reaction time was recorded, participants' responses were monitored to ensure compliance with instructions.

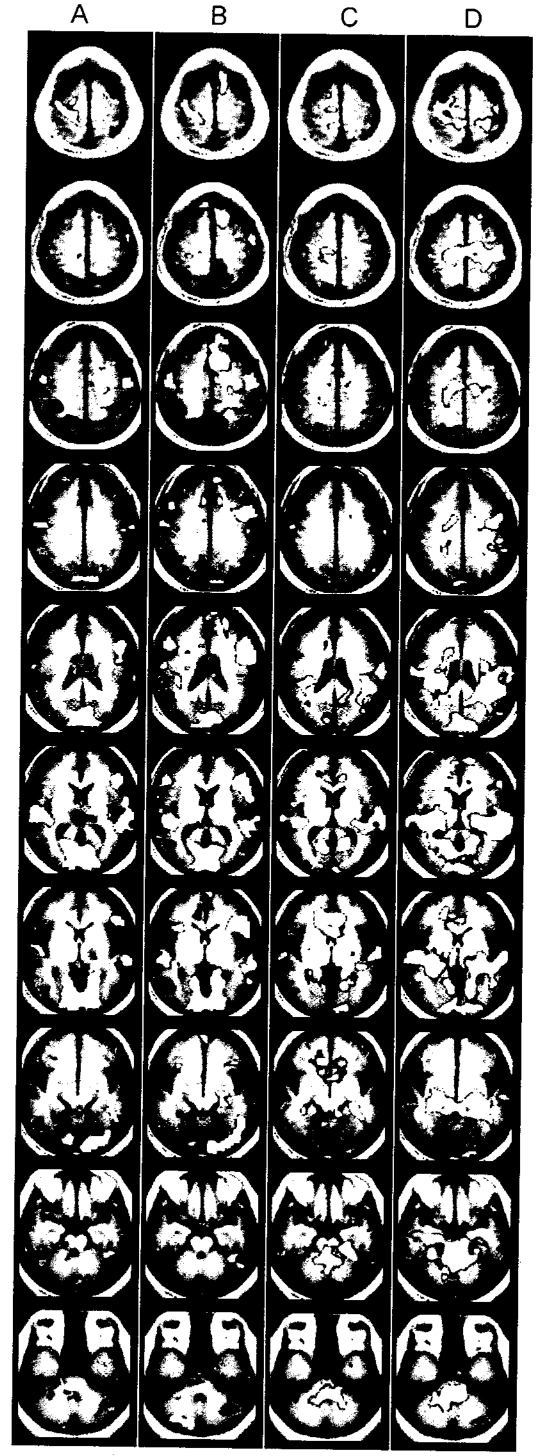

Fig. 3. From left to right: (A) lexical decision repeated minus baseline; (B) lexical decision new minus baseline; (C) naming repeated minus baseline; (D) naming new minus baseline. Increases in activity are colored red/yellow; decreases are colored purple/blue.

The primary results from this experiment are shown as activation maps in Figs. 3 and 4. In Fig. 3, the four ba-

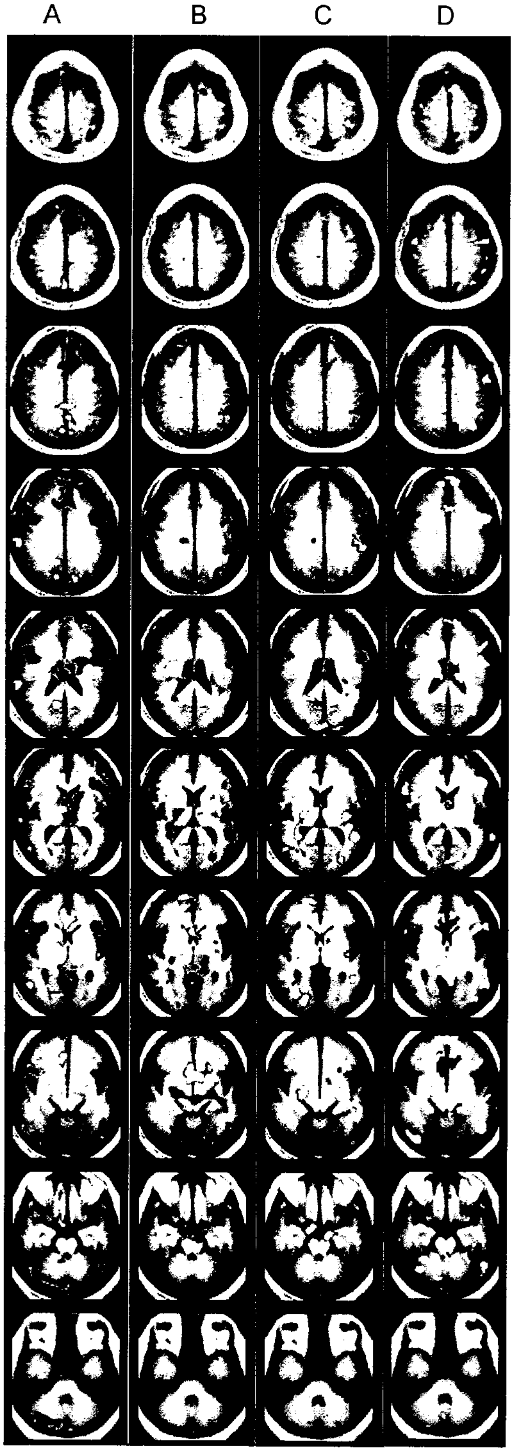

Fig. 4. From left to right: (A) comparison of repeated vs. new words in lexical decision; (B) comparison of repeated vs. new words in naming; (C) differences in change due to training, between tasks; (D) similarities in change due to training, both tasks. Increases in activity are colored red/yellow; decreases are colored purple/blue.

sic maps are shown, in four columns, left to right: lexical decision repeated minus baseline; lexical decision new minus baseline; naming repeated minus baseline; naming new minus baseline. A large degree of overlap between the tasks is evident in the baseline maps; robust activation across major reading-related areas is seen. Some caution is necessary in interpreting these maps. Recall that the baseline conditions were different for lexical decision and naming; each baseline condition was tailored to eliminate or reduce artifacts associated with its respective task. Although the difference in baseline tasks requires caution in making comparisons between tasks, comparisons due to repetition are within-task; there, baseline differences are less problematic.

Fig. 4 shows for lexical decision (Column A) and naming (Column B) those cortical regions that showed reliable training-related changes in activation (i.e., the difference between blocks of repeated and new items). In Column C, regions that showed a *difference* in training-related changes between the two tasks are shown (i.e., it shows those regions that differ reliably between Columns A and B). In Column D, are areas that showed *similar* training-related changes for both tasks (i.e., an intersect analysis; cf. Hadjikhani & Roland, 1998). Note that task invariant changes were almost exclusively reductions. However, in Columns A and B some areas show increases with repetition; these are more prevalent in lexical decision than in naming.

Importantly with respect to our primary hypothesis, the repetition reduction effect at IFG was much stronger for lexical decision than for naming although both tasks showed decreases. These reductions were bi-hemispheric. The relevant sites, shown in Column C, are located primarily in pars opercularis (BA 44) and overlap with sites implicated in the regularity effect in previous studies (Fiez et al., 1999; Herbster et al., 1997; Poldrack & Wagner, 2004; Pugh et al., 1997). For example, individuals who activated IFG bilaterally more also showed stronger regularity effects in lexical decision in Pugh et al. Thus, we may tentatively infer that the greater repetition-induced reduction in the regularity effect in lexical decision (in Experiment 1), compared to naming (in Experiment 2), can be associated, at the neurobiological level of analysis, with the much greater reduction in IFG activity in lexical decision compared to naming. It should be noted, however, that even in the naming task this region did show some reduction with repetition in Experiment 3, presumably because processing became more efficient. Although no significant reduction in the regularity effect was observed in Experiment 2, such a reduction with repetition was found by Visser and Besner (2001). In support of Visser and Besner, recent naming studies (unpublished) in our own laboratory also found reductions in the regularity effect after five repetitions. Nevertheless, comparing Experiments 1, 2, and 3, it is notable that the regularity effect essentially followed the same pattern as IFG activity in lexical decision and naming.

Another difference between tasks was seen in medial-to-right cerebellum where reductions occurred only for lexical decision; no changes were observed in naming. Additionally, at the supplementary motor area, reductions were greater in lexical decision than in naming, mirroring the results obtained in IFG. Thus, repetition reduced activity in regions associated with the control of speech production more in lexical decision than in naming.

Of particular interest is the left hemisphere fusiform gyrus, a region implicated in the development of fluent reading (Shaywitz et al., 2002). The posterior fusiform is associated with pre-semantic "word form" processing, possibly cognitively orthographic (Huang et al., 2001; Tarkiainen et al., 1999). Here, and in other aspects of the ventral pathway, there was more activity for lexical decision than for naming, presumably because the ventral is relatively more engaged in lexical decision. Given that a printed stimulus must generate the same initial activation whether the task is lexical decision or naming, we speculate that the higher activity for lexical decision observed in the word form area was the result of greater feedback, perhaps due to a heightened demand for processing orthographic patterns.

Importantly, there was also a greater degree of repetition-related reduction in this area for lexical decision than for naming. This strong reduction in activity for repeated items in lexical decision suggests that the region required less feedback as items became familiar. In contrast, in the naming task, an emphasis on phonological assembly may make for less involvement of the orthographic processing circuit and, therefore, posterior fusiform did not show as much of a training-related decrease as it did for lexical decision. Note, as can be seen in Column D, that both tasks did show some training-related reduction within this occipitotemporal region but, again, the reductions differed in degree.

Where naming did show greater training-related reductions than lexical decision was at some temporoparietal sites, most notably at more inferior aspects of the supramarginal gyrus and posterior aspects of the superior temporal gyrus (STG) in the LH. In contrast, in the more superior aspect of SMG, similar training-related decreases occurred for both tasks (see Column D, slice 4). We initially expected to find that changes in SMG activity would exactly parallel changes in IFG, i.e., greater reduction in naming over lexical decision. In related work (Mencl et al., submitted for publication), evidence suggested that superior SMG processed phonologically coded information. The present evidence is still consistent with that characterization of SMG but why it did not maintain relatively greater activation under repetition in naming than in lexical decision (as did IFG) is not understood. This functional dissociation between SMG and IFG suggests the need to further explore the similarities and dissimilarities in processing at these two apparently phonological systems.

Greater training-related reductions in activation for naming also occurred at several subcortical and mesiotemporal sites as well as several posterior temporal regions. In particular bi-lateral basal ganglia and thalamus showed reduction for repeated items in the naming task. Again, we can see in the conjunction analysis (Fig. 4) that subcortical reduction for repeated items occurred in both tasks but there was a difference in the degree of reduction.

Finally, we found that for both repeated and new items in both tasks, activation within more anterior aspects of the ventral circuit, particularly the MTG extending into the inferior aspects of STG, activation was constant and stable (see Fig. 3) in both tasks (if anything, a slight increase was observed at some voxels for repeated items in lexical decision). These regions have been implicated in lexical-semantic word processing (Sandak et al., 2004). Thus, in both tasks, the signature of printed word familiarity appears to be reduced activation of occipitotemporal, temporoparietal (i.e., SMG), and inferior frontal areas with activation remaining stable in more anterior temporal lobe regions.

4. General discussion

The behavioral experiments demonstrated that repeatedly presenting an irregular word caused its initial processing disadvantage to disappear quickly if the task was lexical decision (Experiment 1) but not if it was naming (Experiment 2). We interpreted the disappearance of the regularity effect in lexical decision as support for the hypothesis that, with repetition, word recognition in that task becomes less and less dominated by subword phonological analysis. That is, for familiar words, the recognition process is less dependent on the phonological assembly of letters, letter clusters, word bodies, and word rimes. A plausible alternative to assembly is a process by which the lexicon is addressed by larger orthographic units (cf. Coltheart et al., 2001; Ziegler & Goswami, 2005).

We considered other interpretations of the data. Because the lexical decision and naming tasks differed in many respects, any of those differences might be viewed as a possible explanation of why the regularity effect faded in lexical decision but not in naming. For example, the tasks differed in the inclusion of nonwords. We addressed this issue in the discussion section of Experiment 2 and showed that the absence of nonwords in our naming task was likely to have reduced the regularity effect in naming rather than encouraged it. A second difference was that, in naming, the blocks of trials were shorter and repetition was more massed than in lexical decision. This difference would likely produce more rapid learning in naming that, again, should have reduced, rather than encouraged, a regularity effect. A third difference concerned the nature of learning in lexical decision. It is possible that subjects were merely learning a discrimination task in which a simple response (yes or no) was learned to each stimulus. The reduction in the regularity effect would then be due to subjects reading the stimuli not as words and nonwords but as cues for a response. Because the process of recall from the lexicon is relatively effortless compared to learning new associations, such a mechanism seems to us to be less plausible than the alternative but present data cannot rule it out. Finally, there is an alternative explanation based on the fact that large neighborhoods facilitate responding. The slower responses for irregular words in Experiments 1 and 2 might be due to the fact that the average neighborhood size for our irregular words was smaller than that of the regular words. However, there were no effects of neighborhood density on response speed that approached significance and statistical adjustment for density failed to change any analysis substantively.

We suggested that the persistence and greater stability of the regularity effect in naming is likely to be due to the requirement to speak aloud. Speaking necessarily requires recoding of the stimulus into phonological information (phonemes, segments, stress, etc.), a precursor to the spoken articulation. By activating some of that information before lexical identification (e.g., by assembly), information is developed at an early stage that not only assists the lexical identification process but can also prime the subsequent process of articulation. Alternatively, the use of assembly may be promoted by activation of IFG itself because of the requirement to speak. In Experiment 3, we observed (by fMRI) reduced IFG activation in both lexical decision and naming with repetition but a much greater reduction for lexical decision than naming, consistent with the idea that word familiarity in lexical decision, but not in naming, is associated with the larger reductions in speech-related processing. Similar evidence of repetition-induced reductions in activity was produced in the lexical decision task for the supplementary motor area and cerebellum, two regions that, along with IFG, have been implicated in the control of speech production. Thus, the neurobiological evidence suggests reduced assembly with increasing experience and, by implication, a relatively greater contribution from memory-based (i.e., addressed) processing in the ventral circuit.

In neither task was activation of the MTG and inferior aspect of STG affected by repetition (Fig. 4). This suggests that these more anterior ventral sites are functionally distinct from the pre-semantic word form area. Because of decreasing temporoparietal activity with repetition, this constancy in MTG and STG can be viewed as a relatively increased dominance for the ventral circuit and a heightened role for reading familiar items. This shift in relative activity from temporoparietal and anterior to ventral regions possibly underlies some of the general latency reductions observed with repetition in both tasks.

In summary, these behavioral and neuroimaging studies reveal both invariant and specific effects of stimulus repetition in the naming and lexical decision tasks. Task invariant effects included overall reduced decision latencies with concomitant reductions in activation at several major components of the reading circuit (IFG, SMG, and posterior occipitotemporal), along with a constant level of activation maintained at more anterior ventral sites. Specific differences in the behavioral tasks were found in the degree of regularity effect reduction due to repetition (greater reduction in lexical decision than in naming). Parallel effects in brain activation showed relatively greater reductions for lexical decision at IFG, a site previously implicated in the regularity effect and putatively involved in phonological assembly, and at several other regions known to support phonological processing for speech output.

Acknowledgements

This research was supported by a grant to Haskins Laboratories (NICHHD, HD-0194). All experiments reported here were approved by the human subjects Internal Review Boards of Yale University and the University of Connecticut and are in conformance with international standards. All participants gave their informed consent prior to inclusion and were free to withdraw at anytime. We thank Stephen Katz for programming, data collection and analysis of the behavioral experiments. Annette Jenner and Donald Shankweiler offered valuable advice.

Appendix A

Stimuli (regular/irregular, high/low word frequency) in Experiments 1 and 2

Regular high	Regular low	Irregular high	Irregular low
beach	flop	phase	plow
nine	hike	wash	chasm
race	wade	choose	wasp
desk	rust	post	sew
fat	brake	lose	pear
heat	test	touch	slot
drink	pest	break	glove
game	grill	watch	swan
main	slick	none	comb
road	bake	move	hymn
land	hen	dead	isle
real	junk	word	bosom
power	wagon	live	worm
feel	greed	gone	deaf
help	stack	heard	pint
money	peel	love	bury
child	dock	says	cough
face	goat	give	tomb
part	heel	both	shoe
place	gloat	put	flood
make	flag	come	palm
back	clay	some	gross

Nonwords (Experiment 1 only)

Regular high	Regular low	Irregular high	Irregular low
fap	fote	sode	brone
pum	hile	slig	chire
seb	jath	soof	floam
pef	kead	sost	shelk
bine	kive	sove	gwill
bint	lant	soaf	leard
boaf	lart	tace	leasy
carm	leel	swad	plove
calt	lome	tain	mough
dake	mive	thit	phash
clor	moze	tilk	preed
dake	pesh	tose	slock
desk	neen	vask	stily
deze	plat	tury	slace
doot	plog	vone	stelk
dort	mish	woid	stroat

(*Continued*)

Regular high	Regular low	Irregular high	Irregular low
durl	rast	walm	theel
feep	rean	wark	tayes
foad	sace	wast	tound
forn	sagn	yade	tross
gake	seaf	yooly	walch
gead	flad	binch	yombe

Experiment 3 Stimuli: lexical decision

Repeated words	Repeated nonwords
bake	boaf
cough	dake
curl	flib
desk	garm
doom	hote
greed	houch
hood	kinch
hoof	knipe
rust	mosk
swan	neek
warp	plove
wash	stin

New words

batch	clone	blob	crust
gate	curt	boast	dean
groom	dodge	clump	glut
hunk	dread	colt	jolt
math	frown	ditch	marsh
mint	garb	drum	moat
monk	pulp	hoax	mule
pearl	scum	kite	paste
pinch	slug	lame	posh
rake	stub	shrug	raft
trek	void	toss	scalp
twig	weave	wharf	troop

New nonwords

bine	blafe	blean	brame
chone	clag	closk	darf
farn	fleve	flig	flot
gace	gell	goam	grat
hile	jace	jaste	kack
kose	lebe	lenk	lonk
mive	moke	morp	neam
penk	plag	plock	plome
sace	seafe	seft	shenk
sough	spofe	sprate	stam
targe	theel	thorf	tilk
tross	wase	wime	wost

Naming	
Repeated words	
bite	grill
bold	hike
cart	pest
comb	post
couch	probe

(*Continued*)

Naming	
Repeated words	
crow	puff
deaf	rump
dock	slot
float	tile
flop	vase
goat	wade
grape	wasp

New words

bleak	bind	bland	bulk
blunt	cheat	brace	burn
brat	coin	brute	cake
cove	dart	clasp	cheek
crush	doll	crane	cult
dame	flea	curd	dome
drown	fling	dish	draft
folk	glare	drill	drug
freak	goose	feast	flame
gash	gross	forge	flood
graph	lens	frost	fond
jade	loop	gram	fork
ledge	moth	haunt	grab
lick	nest	lard	leaf
mace	peep	loaf	lease
musk	prep	muck	pile
norm	roam	mute	plate
pork	sash	quiz	rent
ramp	shack	reap	silk
slab	smug	sage	skirt
spit	stunt	shawl	vent
thorn	swear	slate	wart
turf	thump	star	wipe
vile	wool	valve	wrist

References

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

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology*, *26*, 439–454.

Carello, C., Turvey, M. T., & Lukatela, G. (1992). Can theories of word recognition remain stubbornly nonphonological? In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning. Advances in psychology: Vol. 94* (pp. 211–226). Amsterdam: Elsevier North Holland Press.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, *125*, 1054–1069.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, *25*, 257–271.

Coltheart, M., Davelaar, E., Jonasson, J. F., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & performance: Vol. VI* (pp. 535–555). Erlbaum: Hillsdale, NJ.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1197–1211.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Fiebach, C. J., Friederici, A. D., Mueller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*(1), 11–23.

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading that words are pronounced more rapidly than nonwords. *Neuron*, *24*, 205–218.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, *2*, 165–189.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 104–115.

Hadjikhani, N., & Roland, P. E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: A positron emission tomographic study. *Journal of Neuroscience*, *18*, 1072–1084.

Harm, M., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491–528.

Harm, M., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*, 155–182.

Helenius, P., Tarkiainen, A., Cornelissen, P., Hansen, P. C., & Salmelin, R. (1999). Dissociation of normal feature analysis and deficient processing of letter-strings in dyslexic adults. *Cerebral Cortex*, *4*, 476–483.

Helenius, P., Uutela, K., & Hari, R. (1999). Auditory stream segregation in dyslexic adults. *Brain*, *122*, 907–913.

Herbster, A., Mintun, M., Nebes, R., & Becker, J. (1997). Regional cerebral blood flow during word and nonword reading. *Human Brain Mapping*, *5*, 84–92.

Herdman, C. M., Chernecki, D., & Norris, D. (1999). Naming cAsE aLtErNaTeD words. *Memory & Cognition*, *27*(2), 254–266.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 166–183.

Huang, J., Carr, T. H., & Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping*, *15*, 39–53.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory & Language*, *36*, 505–529.

Kircher, T. T. J., Brammer, M. J., Levelt, W., Bartels, M., & McGuire, P. K. (2004). Pausing for thought: Engagement of left temporal cortex during pauses in speech. *NeuroImage*, *21*, 84–90.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., & Buckner, R. L. (2000). Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage*, *11*, 735–759.

Mencl, W. E., Frost, S., Sandak, R., Mason, S., Lee, J. R., Rueckl. J., Constable, T., et al. Effects of orthographic and phonologic priming in printed word identification: An fMRI study, submitted for publication.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, M. (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.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *NeuroImage*, *14*, 182–193.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *101*, 56–115.

Poldrack, R. A., & Wagner, A. D. (2004). What can neuroimaging tell us about the mind? *Current Directions in Psychological Science*, *13*, 177–181.

Price, C. J., Winterburn, D., Giraud, A. L., Moore, C. J., & Noppeney, U. (2003). Cortical localization of the visual and auditory word form areas: A reconsideration of the evidence. *Brain and Language*, *86*, 272–286.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. A., Shankweiler, D. P., Katz, L., Fletcher, J. M., et al. (1997). Predicting reading performance from neuroimaging profiles: The cerebral basis of phonological effects in printed word identification. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 1–20.

Pugh, K. R., Mencl, E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Skudlarski, P., et al. (2000). The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity in posterior cortex. *Psychological Science*, *11*, 51–56.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Review*, *6*, 207–213.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Lee, J. R., Katz, L., Frost, S. J., et al. (2001). Neuroimaging studies of reading development and reading disability. *Learning Disabilities Research and Practice*, *16*, 240–249.

Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., et al. (2001). Neurobiological studies of reading and reading disability. *Journal of Communication Disorders*, *34*, 479–492.

Rumsey, J. M., Nace, K., Donohue, B., Wise, D., Maisog, J. M., & Andreason, P. (1997). A positron emission tomographic study of impaired word recognition and phonological processing in dyslexic men. *Archives of Neurology*, *54*, 562–573.

Sandak, R., Mencl, W. E., Frost, S. J., Rueckl, J., Katz, L., Moore, D., et al. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective, and Behavioral Neuroscience*, *4*, 67–88.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 1–17.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Shaywitz, S. E., Shaywitz, B. A., Pugh, K., Fulbright, R., Constable, R. T., Mencl, W. E., et al. (1998). Functional disruption in the brain for reading in dyslexia. *Proceedings of the National Academy of Science of the United States of America*, *95*, 2636–2641.

Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., et al. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, *52*, 101–110.

Seidenberg, M. S. (1992). Beyond orthographic depth in reading: Equitable division of labor. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 85–118). Amsterdam: Elsevier.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tannenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning & Verbal Behavior*, *23*, 383–404.

Skudlarski, P., Constable, R., & Gore, J. C. (1999). ROC analysis of statistical methods used in functional MRI: Individual subjects. *NeuroImage*, *9*, 311–329.

Tabachnick, B. G., & Fidell, L. S. (2001). *Computer-assisted research design and analysis*. Boston: Allyn and Bacon.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation of words in two writing systems: Italian and English. *Memory and Cognition*, *20*, 303–313.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. M. Rayport, Trans.* New York: Thieme Medical Publishers.

Tarkiainen, A, Helenius, P., Hansen, P., Cornelissen, P., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain*, *122*, 2119–2132.

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

Visser, T., & Besner, D. (2001). On the dominance of whole-word knowledge in reading aloud. *Psychonomic Bulletin and Review*, *8*, 560–567.

Xu, B., Grafman, J., Gaillard, W. D., Ishii, K., Vega-Bermudez, F., Pietrini, P., et al. (2001). Conjoint and extended neural networks for the computation of speech codes: The neural basis of selective impairment in reading words and pseudowords. *Cerebral Cortex*, *11*, 267–277.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain-size theory. *Psychological Bulletin*, *131*, 3–29.