Effects of Lexicality, Frequency, and Spelling-to-Sound Consistency on the Functional Anatomy of Reading

Julie A. Fiez,*§ David A. Balota,‡
Marcus E. Raichle,† and Steven E. Petersen†‡
*Departments of Psychology and Neuroscience and
Center for the Neural Basis of Cognition
University of Pittsburgh
Pittsburgh, Pennsylvania 15260
†Department of Neurology and Neurological Surgery,
Department of Anatomy and Neurobiology, and
Department of Radiology
Washington University School of Medicine
‡Department of Psychology
Washington University
St. Louis, Missouri 63110

Summary

Functional neuroimaging was used to investigate three factors that affect reading performance: first, whether a stimulus is a word or pronounceable nonword (lexicality), second, how often a word is encountered (frequency), and third, whether the pronunciation has a predictable spelling-to-sound correspondence (consistency). Comparisons between word naming (reading) and visual fixation scans revealed stimulus-related activation differences in seven regions. A left frontal region showed effects of consistency and lexicality, indicating a role in orthographic to phonological transformation. Motor cortex showed an effect of consistency bilaterally, suggesting that motoric processes beyond high-level representations of word phonology influence reading performance. Implications for the integration of these results into theoretical models of word reading are discussed.

Introduction

The ability to transform orthographic (letter and letter sequence) information into phonological (speech sound) information underlies our ability to read words aloud, and it is one of the most important skills we learn. In this study, we use neuroimaging to examine the neural substrates involved in orthographic to phonological transformation. The experimental design manipulates three factors that affect reading performance: first, whether a stimulus is a word or a nonword (lexicality), second, how often the stimulus is encountered (frequency), and third, whether the pronunciation of the word is predictable based upon its spelling (spelling-to-sound consistency). We hypothesized that these factors could serve as markers for brain regions involved in different aspects of orthographic to phonological transformation because they are thought to influence behavior for different reasons.

Considering first the influence of lexicality, it is known

that words are pronounced more rapidly than nonwords (Forster and Chambers, 1973; Glushko, 1979; Monsell et al., 1992). Repeated exposure to an item may lead to the development of stored word-specific (localist) orthographic, phonological, and semantic representations that permit rapid and efficient transformations (Forster and Chambers, 1973; Morton and Patterson, 1980; Coltheart et al., 1993; Besner, 1999). Alternatively, repeated exposure may enhance the strength of distributed and interactive connections between orthographic, phonological, and semantic representations (Seidenberg and McClelland, 1989; Plaut et al., 1996). In the absence of prior experience, pronunciation may involve slower alternative or additional processes, such as a rule-based system that transforms sublexical orthographic to sublexical phonological units (e.g., letters or letter clusters to phonemes) in order to assemble a pronunciation by "sounding out" the nonword (Meyer et al., 1974; Morton and Patterson, 1980; Coltheart et al., 1993; Besner, 1999). Consideration of these issues led us to hypothesize that effects of lexicality could serve as markers for either word-specific representations of orthography, phonology, or semantics, or for processes engaged in the absence of preexisting lexical representations.

Turning to effects of word frequency, it is known that frequently occurring words are pronounced more rapidly than words that occur less often (Forster and Chambers, 1973; Frederiksen and Kroll, 1976; Scarborough et al., 1977; Balota and Chumbley, 1984; Pugh et al., 1997). The ubiquitous nature of frequency effects across a wide range of other tasks indicates that experience with a word may enhance many different types of processing (Balota and Chumbley, 1985; Jescheniak and Levelt, 1994). By itself, an effect of whole word frequency thus reveals little about the type of information represented (e.g., orthographic versus phonological), but it does reveal something about the level of representation. Effects of word frequency may serve as markers for regions involved specifically in the access and representation of either localist or distributed lexical level information.

Finally, considering the effects of consistency, it is critical that orthographic to phonological transformations in English are only quasiregular (Plaut et al., 1996): while visually similar words tend to have similar pronunciations (e.g., *hint*, *mint*, *lint*), exceptions occur (e.g., *pint*). Words that follow the normal patterns of spelling-to-sound correspondence are pronounced more rapidly than words that are exceptions to these patterns (Andrews, 1982; Monsell et al., 1992; Jared, 1997). Tasks that explicitly require orthographic to phonological transformation (e.g., reading aloud a single word) usually produce robust behavioral effects of spelling-to-sound consistency. Tasks that do not require such transformation, such as lexical decision tasks, usually produce smaller effects (Andrews, 1982; Pugh et al., 1997; Taft and van Graan, 1998). In terms of functional brain activation, we hypothesized that spelling-to-sound consistency effects could serve as markers for brain regions

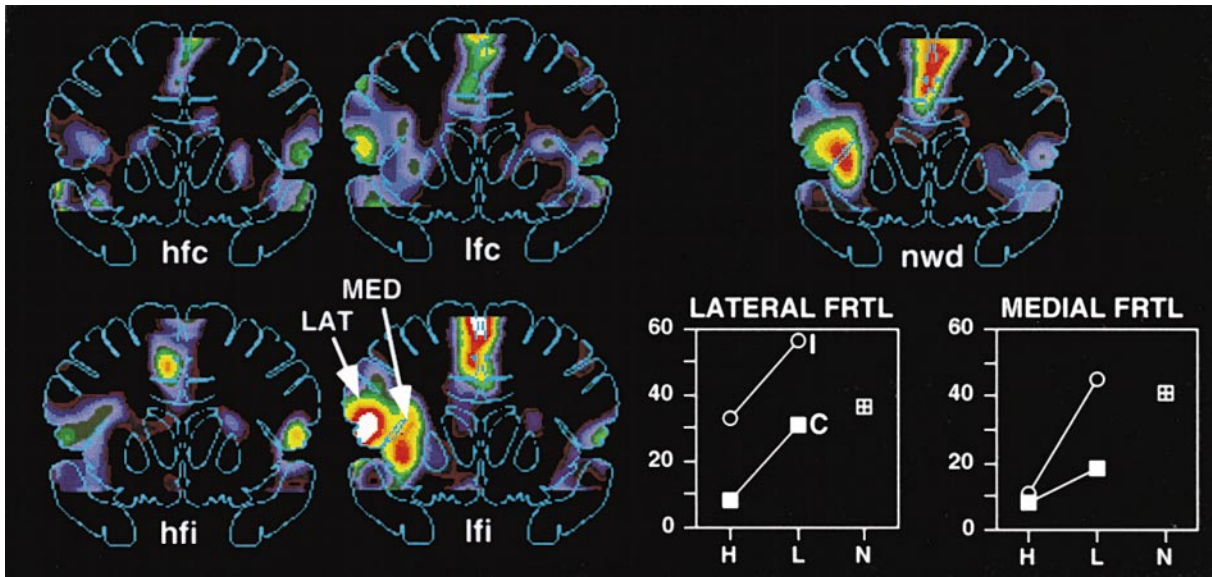§To whom correspondence should be addressed (e-mail: fiez+@pitt.edu).

Figure 1. Left Lateral and Medial Frontal Opercular Activation Is Shown across the Four Word Conditions and the Pronounceable Nonword Condition

The images are coronal sections located 13 mm anterior to the anterior commissure, with magnitudes represented by a color scale that ranges from 0 (dark purple) to 60 (white) counts. The graphs show mean regional magnitudes across frequency (H, high; L, low; N, nonword), with inconsistent words indicated by the open circles, consistent words indicated by the closed squares, and nonwords indicated by the patterned square. hfc, high-frequency consistent; lfc, low-frequency consistent; hfi, high-frequency inconsistent; lfi, low-frequency inconsistent; nwd, nonword condition.

involved specifically in the process of orthographic to phonological transformation.

Behaviorally, effects of consistency interact with effects of frequency. Subjects are particularly slow to begin reading low-frequency inconsistent words (e.g., *pint*), as compared to low-frequency consistent words (e.g., *mint*), high-frequency inconsistent words (e.g., *have*), and high-frequency consistent words (e.g., *gave*) (Andrews, 1982; Seidenberg et al., 1984; Taraban and McClelland, 1987; Monsell et al., 1992). This interaction is thought to reflect a trade-off between two capacities that skilled readers often take for granted: they can use information about consistency to make generalizations that support the pronunciation of novel words (e.g., nonwords such as *sint*), but they can also accurately pronounce words that violate these generalizations (e.g., *pint*).

There has been extensive debate in the cognitive psychological literature about whether the pronunciations of nonwords versus inconsistent words are supported by distinct computational processes, or whether they are supported by a single computational mechanism. In dual route frameworks, the two capacities are supported by two different processing routes for orthographic to phonological transformation. For instance, in the model proposed by Coltheart and colleagues (Coltheart et al., 1993) a direct route translates entire word form representations into an articulatory output using an associative, frequency-dependent process. An assembled route allows a word or nonword to be "sounded out" by translating individual letter units into corresponding sound units using a rule-based process. The interaction between frequency and consistency arises from competition between the two routes (Marshall and Newcombe,

1973; Morton and Patterson, 1980; Coltheart et al., 1983; Besner and Smith, 1992; Monsell et al., 1992; Coltheart et al., 1993).

The development of connectionist networks trained to pronounce visually presented words and nonwords called into question the basic assumptions of the dual route framework (Seidenberg and McClelland, 1989; Plaut et al., 1996). The networks demonstrate frequency and consistency effects similar to those found in normal reading, even though the computations are performed using a parallel distributed process without any explicit coding of the spelling-to-sound rules of English. These computational results form the basis for an alternative theoretical framework, in which frequency and consistency effects are thought to arise through a single process that is sensitive to statistical relationships between distributed representations that develop through repeated word exposure (cf., Glushko, 1979; Marcel, 1980). The same set of processing components and the same computational algorithm drive correct pronunciations for both inconsistent words and nonwords.

There are three principle objectives of this study. The first is to draw distinctions between the regions that collectively interact to support orthographic to phonological transformation. We hypothesize that the brain regions that are active during a word reading task will show different patterns of activation in response to stimulus factors that serve as markers for different computational processes. The second is to use neuroimaging data to provide new insights into the cognitive processes and brain regions that support orthographic to phonological transformation. We hypothesize that some regions will show a pattern of activation that does not track with behavioral performance. Attempts to interpret
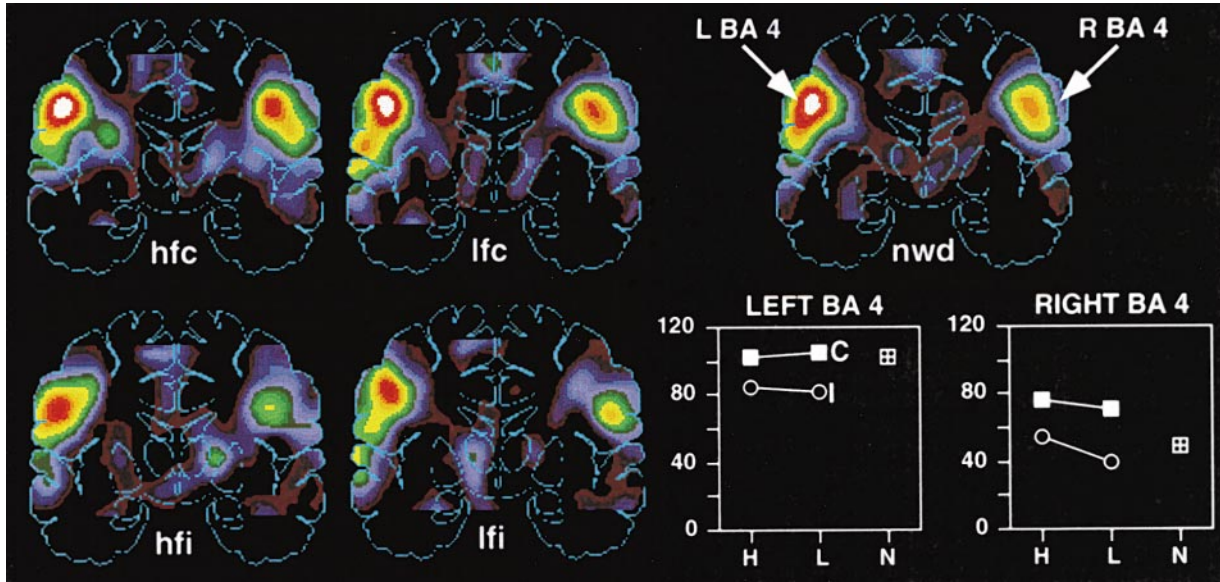
Figure 2. Left and Right Precentral Gyral Activation Is Shown across the Four Word Conditions and the Pronounceable Nonword Condition

The images are coronal sections (BA 4) located 9 mm posterior to the anterior commissure, with magnitudes represented by a color scale that ranges from 0 (dark purple) to 100 (white) counts. The graphs show mean regional magnitudes across frequency (H, high; L, low; N, nonword), with inconsistent words indicated by the open circles, consistent words indicated by the closed squares, and nonwords indicated by the patterned square. hfc, high-frequency consistent; lfc, low-frequency consistent; hfi, high-frequency inconsistent; lfi, low-frequency inconsistent; nwd, nonword condition.

the functional role of these areas will require a new focus on mechanisms that could give rise to dissociations between frequency, consistency, and lexicality effects, rather than the interactions observed in behavioral performance. The third is to examine the relationship between neuroimaging results and dual route versus connectionist models of reading. Attempts to "map" brain regions onto components of these models may fail, because the models have not been developed to account for regional dissociations between stimulus factors, or patterns of converging activation observed across reading and other phonological tasks.

To test our hypotheses, subjects will be asked to read aloud five types of items, in separate blocks: (1) high-frequency consistent words, (2) high-frequency inconsistent words, (3) low-frequency consistent words, (4) low-frequency inconsistent words, and (5) pronounceable nonwords. Functional brain activation will be measured during each block, using positron emission tomography (PET). Our analyses will first identify brain regions with more blood flow during reading than visual fixation and then identify which of these regions show effects of frequency, consistency, or lexicality.

Results

The results from this study are presented in several sections. The first section describes the behavioral results. The second section provides an overview of the functional imaging data set. The third section describes reliable regions of activation associated with word reading. The fourth section describes the effects of frequency and consistency on the activation of regions associated with word reading. The fifth section describes the effects of lexicality on the same set of regions, and it describes areas of activation uniquely identified in the nonword condition. The sixth section further assesses word versus nonword differences through an analysis of previously collected data from a verbal working memory study.

For simplicity, we will use the term consistency throughout the manuscript, though it should be noted that in dual route versus connnectionist frameworks there is a distinction between regularity and consistency. In dual route models, the notion of regularity is binary, with regular words defined as those whose pronunciation follows grapheme to phoneme translation

Table 1. Naming Latencies and Error Rates across Stimulus Conditions

Stimulus Condition	Latency (Mean ms ± SD)		Errors (Mean % ± SD)		
	Correct	Incorrect	Atypical	Dysfluent	Other
high-frequency consistent	468 ± 48	470 ± 46	0 ± 0	0.2 ± 0.6	0.4 ± 0.8
high-frequency inconsistent	463 ± 53	465 ± 56	0.4 ± 0.8	0.2 ± 0.6	0.4 ± 1.2
low-frequency consistent	475 ± 46	476 ± 46	0 ± 0	0.5 ± 1.3	0.9 ± 2.4
low-frequency inconsistent	547 ± 48	562 ± 54	8.2 ± 5.3	1.8 ± 2.1	7.3 ± 3.3
pronounceable nonwords	538 ± 64	546 ± 64	1.8 ± 0.9	0.8 ± 1.0	7.2 ± 4.6

Table 2. Positive Regional Activation across Hypothesis Generation and Test Groups

Region Location in Test Group				Statistics in Text Group				Overall Crds		
Gyrus (BA)	X	Y	Z	Mag	Df	t Val	p Val	X	Y	Z
Significant/near significant changes										
L inferior frtl (BA 44)	−51	9	8	26	10	3.63	0.0023***	−49	11	10
R inferior frtl (BA 44)	57	19	10	11	10	1.38	0.0996*	53	15	8
superior frtl (BA 6)	−1	5	54	43	7	3.32	0.0064*	−1	5	56
L precentral (BA 4)	−47	−13	34	92	10	11.58	<0.0001***	−47	−13	34
R precentral (BA 4)	49	−5	24	44	10	3.97	0.0013***	49	−7	26
L middle insula	−29	−3	24	16	10	1.61	0.0692*	−27	−5	22
L sup temp (BA 42)	−47	−31	12	23	10	2.11	0.0306*	−49	−33	10
R sup temp (BA 42)	53	−29	8	39	10	2.95	0.0073*	55	−29	8
L fusiform (BA 37)	−43	−45	−8	30	10	4.37	0.0007***	—	—	—
R fusiform (BA 37)	43	−49	−10	16	10	1.93	0.0414*	33	−37	−10
L fusiform (BA 19)	−43	−67	−4	45	10	3.89	0.0015***	−43	−63	−4
L putamen	−13	−31	12	18	10	2.01	0.0364*	−19	−33	16
medial cerebellum	−3	−69	−12	42	10	3.91	0.0015***	1	−71	−16
L paramedial cblm	−19	−67	−16	54	7	4.65	0.0012***	−17	−67	−16
R paramedial cblm	15	−69	−12	53	10	4.88	0.0003***	17	−67	−14
R lateral cerebellum	33	−69	−12	28	10	2.02	0.0358*	33	−67	−12
Nonsignificant changes										
R lateral cerebellum	49	−65	−16	12	7	1.09	0.1561			

For each region identified in a hypothesis generation data set, regional magnitudes in a test data set were analyzed using a one-sample t test. For each region of significant (***p < 0.05 following Bonferroni correction for 17 comparisons) or near significant (*p < 0.10 prior to Bonferroni correction) change, the nearest focus in the overall image that combined both data sets was identified. For the focus at or near L BA 37, a focus separate from the more posterior activation at or near BA 19 could not be identified in the overall image.
BA, Brodmann area; Mag, magnitude; Df, degrees of freedom; t Val, t values; p Val, p values.

"rules." In connectionist models, the term consistency is used to capture the idea that regularity is graded; for example, there are items (e.g., *good*) whose pronunciations do not follow the rules, but which are nonetheless consistent with many other orthographically similar words (Plaut et al., 1996).

Behavioral Results
Table 1 summarizes the mean pronunciation latencies and error rates across the different conditions. The data were analyzed using a two-way analysis of variance (ANOVA), with frequency and consistency as within-subject factors. As expected, for both total error rate and pronunciation latency, there were significant (p < 0.001) main effects of frequency ($F[1,10] = 72.26$ for total errors, $F[1,10] = 54.64$ for pronunciation latency) and consistency ($F[1,10] = 63.04$ for total errors, $F[1,10] = 35.19$ for pronunciation latency). The expected interaction between frequency and consistency reached significance for both errors and pronunciation latency ($F[1,10] = 67.36$, p < 0.0001, and $F[1,10] = 17.08$, p = 0.002, respectively). The results for nonwords were also typical: the error rate and pronunciation latency were larger than the values obtained high-frequency consistent, high-frequency exception, and low-frequency consistent words, but were similar to those obtained for low-frequency inconsistent words. While incorrect responses were excluded from the latency analysis reported above, the same pattern of results were found when they were included.

Imaging Results: Overview of the Data Set
The interpolated and anatomically transformed images from each subject were used to create individual difference images of a single reading task scan minus a fixation control scan from the same subject. For word reading, the data set consisted of 44 read minus fixation individual difference images (four read conditions from each of 11 subjects). For nonword reading, the data set consisted of 20 nonword minus fixation individual difference images: two repetitions of the nonword condition from each of nine subjects, and one repetition from two other subjects (data were lost because of technical difficulties and excessive movement). Individual difference images were averaged to create mean difference images; for instance, an overall word reading minus fixation mean difference image was created by averaging the 44 word reading minus fixation individual difference images. Mean difference images were used to identify foci of change during a particular task condition. These foci defined the centers of spherical regions of interest that were applied to the individual difference images in order to generate individual regional magnitudes for statistical evaluation.

Region centers are referenced through three coordinate distances (x, y, z; in mm) in the space of the Talairach and Tournoux (Talairach and Tournoux, 1988) stereotactic atlas: the distance to the right (+) or left (−) of midline, the distance anterior (+) or posterior (−) to the anterior commissure, and the distance above (+) or below (−) the horizontal plane through the anterior and posterior commissures. Activation changes are in normalized counts (e.g., a peak in a difference image with a magnitude of +100 counts represents very approximately a 10% increase in activity or blood flow).

Analysis of Word Reading: Regions of Significant and Reliable Change Replication across Data Sets and Task Conditions
We used a two-stage approach to identify candidate regions of interest associated with word reading and then to test which of these regions were significantly active in an independent data set (cf., Buckner et al.,

Table 3. Negative Regional Activation across Hypothesis Generation and Test Groups

Region Location in Test Group				Statistics in Test Group				Overall Crds		
Gyrus (BA)	X	Y	Z	Mag	Df	t Val	p Val	X	Y	Z
Significant/near significant changes										
R middle frtl (BA 10)	25	53	20	−26	10	−5.09	0.0002***	25	53	14
R orbital frtl (BA 10)	11	51	−7	−14	7	−1.70	0.0669*	13	51	−7
L middle frtl (BA 10/46)	−21	43	14	−25	10	−2.27	0.0234*	−25	47	10
R middle frtl (BA 46)	39	45	14	−13	10	−1.54	0.0768*	41	47	14
R middle frtl (BA 9/46)	31	29	28	−24	10	−2.22	0.0255*	37	27	30
L middle frtl (BA 8)	−39	23	38	−18	10	−2.01	0.0364*	−37	19	40
R middle frtl (BA 8)	31	17	40	−39	10	−6.75	<0.0001***	31	21	38
L middle frtl (BA 6)	−27	1	52	−12	9	−1.89	0.0458*	−27	1	52
R middle frtl (BA 6)	25	−7	52	−11	9	−1.74	0.0584*	27	−5	52
R post cingulate (BA 23)	21	−51	24	−13	10	−2.47	0.0165*	25	−49	32
ant precuneus (BA 31)	3	−43	34	−39	10	−3.37	0.0036*	3	−43	34
post precuneus (BA 31)	3	−67	28	−38	10	−5.05	0.0002***	—	—	—
ant precuneus (BA 7)	−3	−51	53	−24	7	−2.48	0.0212*	5	−47	54
post precuneus (BA 7)	1	−71	44	−46	10	−6.74	<0.0001***	−3	−65	44
L precuneus (BA 7)	−17	−73	36	−30	10	−4.23	0.0009***	−15	−71	36
R sup parietal (BA 7)	15	−63	48	−38	8	−4.16	0.0016***	13	−61	48
R sup parietal (BA 7)	31	−51	50	−17	8	−3.06	0.0078*	39	−49	50
L supramarg (BA 39/40)	−35	−61	32	−28	10	−3.87	0.0015***	−35	−61	36
R supramarg (BA 39/40)	47	−57	38	−23	10	−2.35	0.0203*	51	−55	30
R inf parietal (BA 40)	45	−41	38	−34	10	−4.06	0.0011***	43	−43	42
R occipital (BA 19)	35	−77	26	−22	10	−2.63	0.0125*	31	−71	32
Nonsignificant changes										
middle frontal (BA 10)	−1	65	8	−5	10	−0.57	0.2906			
L ant cingulate (BA 24)	−3	23	18	8	10	0.90	0.8048			
R ant cingulate (BA 24)	7	23	−10	−4	9	−0.50	0.3139			
L ant cingulate (BA 32)	−21	15	40	−1	10	−0.40	0.3506			
R ant cingulate (BA 32)	11	9	42	2	10	0.19	0.5743			
R occipital (BA 18)	−27	−91	24	−6	10	−0.59	0.2839			

Regional magnitudes in the test data set were analyzed using a one-sample t test. For each region of significant (***p < 0.05 following Bonferroni correction for 27 comparisons) or near significant (*p < 0.10 prior to correction) change, the nearest focus in the overall word minus fixation image was identified. The two foci near BA 31 merged into a single focus in the overall image.
BA, Brodmann area; Mag, magnitude; Df, degrees of freedom; t Val, t values; p Val, p values.

1996; Fiez et al., 1996; Hunton et al., 1996; Shulman et al., 1997).

In the first stage, the individual word minus fixaton difference images related to the four word stimulus conditions (four scans from each of 11 subjects) were divided into two groups of 22 images: one for hypothesis generation and the other for hypothesis testing. The groups were balanced for condition and scan order, and (with one exception) there was no overlap in the fixation control images between the two groups. A mean difference image was created from the 22 hypothesis-generating images, and all foci with a magnitude of 25 counts or greater and a descriptive t value of p < 0.05 were identified; the location of each focus passing these criteria defined the center of a spherical candidate region of interest.

In the second stage, the average magnitude within each candidate region of interest was computed for each of the 22 images in the hypothesis-testing group. For each region, the values from each subject were averaged, and the resulting 11 values were submitted to a one-sample one-tailed t test. For each significant or near significant region of change, the coordinates of the nearest focus of change in the overall word reading minus fixation image were identified in order to best estimate the response location.

Eight candidate regions of positive change replicated in the hypothesis-testing group of images (p < 0.05, one-tailed t test, Bonferroni corrected) (see Table 2). Significant changes were found bilaterally at or near Brodmann area (BA) 4 in the precentral gyrus and in the paramedial cerebellum, along with another midline cerebellar focus. Left-lateralized changes were found at or near BA 44 in the lateral inferior frontal gyrus and in the fusiform gyrus (one region near the border of BA 19 and a more anterior region at or near BA 37). Several other regions showed a trend toward positive activation (p < 0.10 prior to Bonferroni correction), including regions at or near BA 6 (supplementary motor cortex), BA 37 on the right, BA 42 bilaterally, the left putamen, left insula, and right lateral cerebellum.

Eight candidate regions of negative change replicated in the hypothesis-testing group of images (p < 0.05, one-tailed t test, Bonferroni corrected) (see Table 3). Two of these regions were located in the right frontal cortex, at or near BAs 8 and 10. Posteriorly, bilateral regions were found in the medial and lateral parietal cortex (at or near BAs 7 and 40) and near the midline in the precuneus (at or near BA 31). Additional regions that approached significance were found in the middle frontal gyrus bilaterally (at or near BAs 6, 8, and 46) and in the right parietal lobe (at or near BAs 7, 31, 23, 39, and 40).

Reliable Regions of Change in Each Condition

A potential problem with the hypothesis generation/test analysis is that an area of activation associated with

Table 4. Reliable Regions of Positive Change across Different Stimulus Conditions

Region	Overall Word Image			Reliable Changes for Each Condition			
Gyrus	Location	Rep	Rel	Hfc	Hfi	Lfc	Lfi
L inf frtl (45)/ant ins	−35, 15, 6	new					−31, 23, 14
L inf frtl (44)	−49, 11, 10	***	√		(−51, 5, 10)	(−53, 19, 14)	−49, 13, 14
R inf frtl (44)	53, 15, 8	*	√	(57, 19, 2)	(51, 15, 12)		(57, 21, 6)
sup frtl (6/SMA)	−1, 5, 56	*	√	(−1, 3, 50)		(−1, 3, 54)	1, 9, 56
L precentral (4)	−47, −13, 34	***	√	−49, −13, 36	−49, −13, 30	−47, −15, 32	−47, −11, 36
R ant precentr (4)	49, −7, 26	***	√	51, −9, 24	(55, −3, 24)	45, −7, 30	49, −9, 24
R post precentr (4)	39, −13, 34	new		41, −13, 36	41, −11, 30		(37, −15, 36)
L middle insula	−27, −5, 22	*	√	(−29, −9, 22)			
R sup temp (42)	55, −29, 8	*	√	53, −25, 8	55, −29, 8	(51, −31, 8)	
L sup temp (22)	−59, −49, 18	new				−59, −47, 18	
L fusiform (19)	−43, −63, −4	***	√	(−43, −63, −4)	−31, −65, −4	−33, −59, −10	−43, −59, 2
medial cerebellum	1, −71, −16	***	√		3, −71, −16		(−1, −69, −16)
ant medial cblm	—	—			−1, −57, −6		
L paramedial cblm	−17, −67, −16	***	√	(−13, −67, −16)		−15, −67, −16	(−17, −67, −16)
R paramedial cblm	17, −67, −14	***		13, −65, −12		(17, −69, −16)	(19, −67, −14)

Listed are all regions of reliable positive change (descriptive $p < 0.0005$) for each word condition (high-frequency consistent [Hfc], high-frequency inconsistent [Hfi], low-frequency consistent [Lfc], and low-frequency inconsistent [Lfi]). The nearest foci of change in the overall word minus fixation image were located (no distinct focus could be found for the new anterior medial cerebellar focus). Regions of significant (***) or near significant (*) change in the hypothesis generation/test analysis, or which reached the reliability threshold (√) across the entire word data set are noted. For comparison across conditions, foci within 10 mm of the overall location that were marginally reliable ($p < 0.01$) are listed in parentheses.

BA, Brodmann area; Rep, replicated; Rel, reliable.

only one condition might be overlooked. For this reason, we also looked for highly reliable regions of change in each condition. Mean reading minus fixation difference images were created for each task condition and also for the overall word reading condition. For each image, all foci of positive and negative change with a corresponding one-sample regional t value equivalent to $p < 0.0005$ were identified as reliable. For each identified region, the nearest focus of change in the overall word reading minus fixation image was identified, in order to obtain an estimate of the region location across conditions.

Positive regions that passed the reliability threshold are listed in Table 4. Three new regions were identified: one in the right precentral gyrus (found in the high-frequency word conditions), one near the border of the insula and left inferior frontal gyrus (found in the low-frequency inconsistent condition), and one in the left superior temporal cortex at or near BA 22 (found in the low-frequency consistent condition). While the same set of reliable regions was not found in every condition, overall the results correspond well with those from the hypothesis generation/test analysis.

The negative regions identified in this secondary analysis correspond well with those described in the first analysis. Only one new region was identified: a ventral frontal region at or near BA 25 ($x = 9$, $y = 11$, $z = -11$), which was found in the low-frequency consistent condition. Further results will not be reported for negative regions; only a few marginally significant stimulus effects were found, and their location in right parietal cortex (BA 7/40) makes it unlikely that they are critically involved orthographic to phonological transformation.

Analysis of Word Reading: Effects of Frequency and Consistency

The first two analyses identified a set of regions associated with word reading. In a third analysis, we investigated the effects of frequency and consistency upon the activation of these regions. We examined the activation in all the word reading minus fixation regions that were significant ($p < 0.05$ following a Bonferroni correction) in the hypothesis generation/test analysis and in all regions that passed the reliability threshold ($p < 0.0005$) in the second analysis. For each region, regional magnitudes were computed for each of the 44 word minus fixation individual difference images and then assessed using an ANOVA with frequency (high versus low) and consistency (consistent versus inconsistent) as within-subject factors.

Table 5 summarizes the regional magnitudes across word conditions. Three regions showed an effect of stimulus consistency (see figures 1 and 2). In the precentral gyrus bilaterally, greater activation was found when subjects read aloud consistent words than inconsistent words, regardless of frequency (on the left: $x = -47$, $y = -13$, $z = 34$; $F[1,10] = 9.1$, $p = 0.013$; and on the right: $x = 39$, $y = -13$, $z = 34$; $F[1,10] = 20.2$, $p = 0.001$). A left lateral inferior frontal gyral region at or near BA 44 also showed an effect of consistency, but in the opposite direction: inconsistent words produced more activation than consistent words ($x = -49$, $y = 11$, $z = 10$; $F[1,10] = 17.2$, $p = 0.002$), with a trend toward lower-frequency words also producing more activation than high-frequency words ($F[1,10] = 4.60$, $p = 0.058$).

Weak effects of word frequency were found in two regions. A left temporal region, located at or near BA 22, was more active when subjects read low-frequency than high-frequency words ($z = -59$, $y = -49$, $z = 18$; $F[1,10] = 5.2$, $p = 0.045$). A region at or near the supplementary motor area (SMA/BA 6) showed a tendency toward greater activation for low-frequency as compared to high-frequency words ($x = -1$, $y = 5$, $z = 56$; $F[1,7] = 3.59$, $p = 0.100$).

Finally, a region located near the border of the inferior frontal gyrus (BA 44/45) and the anterior insula ($x = -35$, $y = 15$, $z = 6$), showed a pattern of activation very

Table 5. Positive Regional Activations across Stimulus Conditions

Region Location				Magnitudes (± SE) by Condition				
Gyrus	X	Y	Z	Hfc	Hfi	Lfc	Lfi	Effect
L inf frtl (44/45)/ant ins	−35	15	6	8 ± 7	11 ± 16	18 ± 6	45 ± 10	F, FxC
L inf frtl (44)	−49	11	10	8 ± 8	33 ± 12	31 ± 8	56 ± 10	C**
R inf frtl (44)	53	15	8	26 ± 9	34 ± 8	26 ± 10	30 ± 11	
sup frtl (6/SMA)	−1	5	56	32 ± 6	37 ± 17	55 ± 13	55 ± 11	
L precentral (4)	−47	−13	34	103 ± 10	85 ± 10	106 ± 11	82 ± 9	C**
R ant precentr (4)	49	−7	26	68 ± 9	46 ± 15	69 ± 12	60 ± 13	
R post precentr (4)	39	−13	34	76 ± 10	54 ± 10	71 ± 12	40 ± 11	C
L middle insula	−27	−5	22	35 ± 13	15 ± 11	21 ± 13	14 ± 9	
R sup temp (42)	55	−29	8	44 ± 8	64 ± 11	46 ± 13	40 ± 17	
L sup temp (22)	−59	−49	18	7 ± 13	12 ± 10	32 ± 7	15 ± 12	F
L fusiform (37)	−43	−45	−8	31 ± 8	28 ± 9	28 ± 8	35 ± 11	
L fusiform (19)	−43	−63	−4	52 ± 14	47 ± 9	38 ± 11	55 ± 10	
medial cerebellum	−1	−57	−6	21 ± 18	55 ± 10	25 ± 14	31 ± 15	
ant medial cblm	1	−71	−16	39 ± 12	68 ± 13	43 ± 18	60 ± 15	
L paramedial cblm	−17	−67	−16	73 ± 19	58 ± 23	71 ± 11	70 ± 15	
R paramedial cblm	17	−67	−14	66 ± 14	45 ± 18	74 ± 12	62 ± 20	

Mean regional magnitude are listed by condition, and effects (p < 0.05) of frequency (F) and consistency (C) are noted. Letters followed by ** indicate significant effects at p < 0.05 following a Bonferroni correction; when the magnitude values were randomly assigned within each region, no significant ANOVA effects (p < 0.05 prior to correction) were found.
BA, Brodmann area; Hfc, high-frequency consistent; Hfi, high-frequency inconsistent; Lfc, low-frequency consistent; Lfi, low-frequency inconsistent.

similar to the behavioral data (see Figure 1). In the low-frequency inconsistent condition, the regional activation was more than twice that found in the other word reading conditions (45 versus less than 20 counts), a result compatible with the fact that the region only surpassed the reliability threshold in the low-frequency inconsistent condition. This frequency by consistency interaction approached significance (F[1,10] = 3.29, p = 0.100), as did a main effect of frequency (F[1,10] = 4.06, p = 0.072). These effects were more reliable (F[1,10] = 5.89, p = 0.036, and F[1,10] = 10.77, p = 0.008, respectively) when a region was placed upon the coordinate location identified in the low-frequency inconsistent minus fixation condition, rather than the nearest peak location in the overall word reading minus fixation image. Effects of lexicality in this region will be reviewed in the next section.

Analysis of Nonword Reading: Lexicality Effects
Lexicality Effects in Regions Identified from Word Conditions

As a first step in comparing the functional anatomy associated with word versus nonword reading, the regions identified during word reading were applied to the individual nonword minus fixation individual difference images. For each subject, the regional magnitude values from the nonword condition were averaged together, as were the four values from the word conditions. Paired t tests were then used to compare the word and nonword values. To compare word/nonword response locations, for each region the focus in the overall nonword minus fixation mean difference image was identified. Finally, for those regions that showed an effect of frequency or consistency in the previous analyses, additional condition-specific word versus nonword comparisons were conducted.

None of the positive regions showed greater activation in the averaged word condition as compared to nonwords (see Table 6). Two regions showed greater

activation for nonwords than words. One region was a left fusiform region located at or near BA 37 (x = −43, y = −45, z = −8; t[1,10] = 2.41, p = 0.04). The second region was located near the border of the insula and the medial inferior frontal gyrus, at or near BA 44/45 (x = −35, y = 15, z = 6; t[1,10] = 2.89, p = 0.02), This same region showed an interaction between frequency and consistency, which led us to conduct a series of condition-specific comparisons as well. Activation in the nonword condition was significantly greater than the high-frequency consistent, high-frequency inconsistent, and low-frequency consistent word conditions (t[1,10] > 2.37, p < 0.05), but not significantly different than the low frequency inconsistent condition (p > 0.05). The overall pattern of activation in this region thus paralled the behavioral data, with the low frequency inconsistent and nonword conditions producing the greatest activation, the slowest reaction time, and the most errors.

Specific comparisons were also conducted for other regions that showed frequency and consistency effects in the word conditions. In the left lateral inferior frontal region (x = −49, y = 11, z = 10) nonwords differed from high-frequency consistent words (t[1,10] = 3.19, p = 0.01), but not the other three word conditions (p > 0.05). Thus, even though this region is within 2 cm of the more medial region, it shows a very different pattern of activation (see also Figure 1). In the left precentral gyrus (BA 4), activation in the nonword condition followed that observed for consistent words: greater activation was found for nonwords versus inconsistent words (t[1,10] = 2.33, p = 0.04), but not for nonwords versus consistent words. In right precentral gyrus, the pattern was less clear: nonwords did not differ from either consistent or inconsistent words. In the left superior temporal region (which showed a frequency effect), the nonword activation did not differ from either the high- or the low-frequency word conditions (p > 0.05). Finally, in SMA the activation in the nonword condition was similar to that

Table 6. Positive Regional Activations for Words versus Pronounceable Nonwords

Gyrus (BA)	Word versus Nonword Crds		Word versus Nonword Magnitude				
	Word Crd	Nwd Crds	Word	Nwd	df	t Val	p Val
Regions from word conditions							
L inf frtl (44/45)/ant ins	−35, 15, 6	−39, 11, 12	20	41	10	2.89	0.02*
L inf frtl (44)	−49, 11, 10	−49, 3, 4	32	36	10	0.53	0.61
R inf frtl (44)	53, 15, 8	49, 17, 10	29	23	10	−0.97	0.36
sup frtl (6/SMA)	−1, 5, 56	−1, 7, 56	45	56	7	1.28	0.24
L precentral (4)	−47, −13, 34	−47, −15, 32	94	103	10	1.24	0.24
R ant precentr (4)	49, −7, 26	47, −7, 30	61	70	10	0.97	0.35
R post precentr (4)	39, −13, 34	—	60	49	10	−0.91	0.38
L middle insula	−27, −5, 22	−15, 7, 18	21	18	10	−0.39	0.70
R sup temp (42)	55, −29, 8	53, −29, 8	48	52	10	0.63	0.54
L sup temp (22)	−59, −49, 18	−53, −49, 20	17	8	10	−1.34	0.21
L fusiform (37)	−43, −45, −8	−41, −45, −6	30	42	10	2.41	0.04*
L fusiform (19)	−43, −63, −4	−37, −65, −8	48	57	10	1.39	0.19
medial cerebellum	−1, −57, −6	7, −59, −10	33	33	10	0.00	1.00
ant medial cblm	1, −71, −16	3, −77, −17	53	49	7	−0.42	0.69
L paramedial cblm	−17, −67, −16	−19, −67, −16	68	61	7	−0.56	0.59
R paramedial cblm	17, −67, −14	17, −67, −14	62	69	7	1.31	0.23
Regions identified in nonword condition							
L sup temp (42)	−49, −33, 10	−57, −41, 6	28	33	10	0.87	0.41
R lateral cblm	33, −67, −14	33, −67, −14	45	71	7	2.79	0.03*

For each region, the coordinates of the nearest foci in the word and nonword conditions were identified. For each focus from the word condition, regional magnitudes in the word and nonword conditions were compared using a paired t test. Differences are noted with an ** for p < 0.05 following a Bonferroni correction for 18 comparisons and * for p < 0.05 prior to correction; when the magnitude values were randomly assigned within each condition, a significant effect (p < 0.05 prior to correction) was found in only one region.
BA, Brodmann area; Nwd, nonword; Df, degrees of freedom; t Val, t values; p Val, p values.

observed for low-frequency words: greater activation was found for nonwords than high-frequency words (t[1,7] = 3.36, p = 0.01), but not low-frequency words (p > 0.05).

Lexicality Effects in Regions Identified from the Nonword Condition

The initial analysis of the nonword data focused only upon regions derived from the word reading data. To determine whether there might be regions of activation unique to nonword reading, the same hypothesis generation/test and reliability analyses applied to the word data were applied to the nonword data. For each region found to be significant in the nonword hypothesis generation/test analysis, or that passed the reliability threshold, the nearest focus in the overall word minus fixation image was identified. For novel regions, paired t tests were performed between the word and nonword conditions.

Nearly all of the identified regions of significant and/or reliable change in the nonword condition corresponded to regions identified during word reading (see Tables 5 and 6). However, two novel regions of positive change were identified. One region was located in the left temporal cortex at or near BA 42 (x = −57, y = −33,
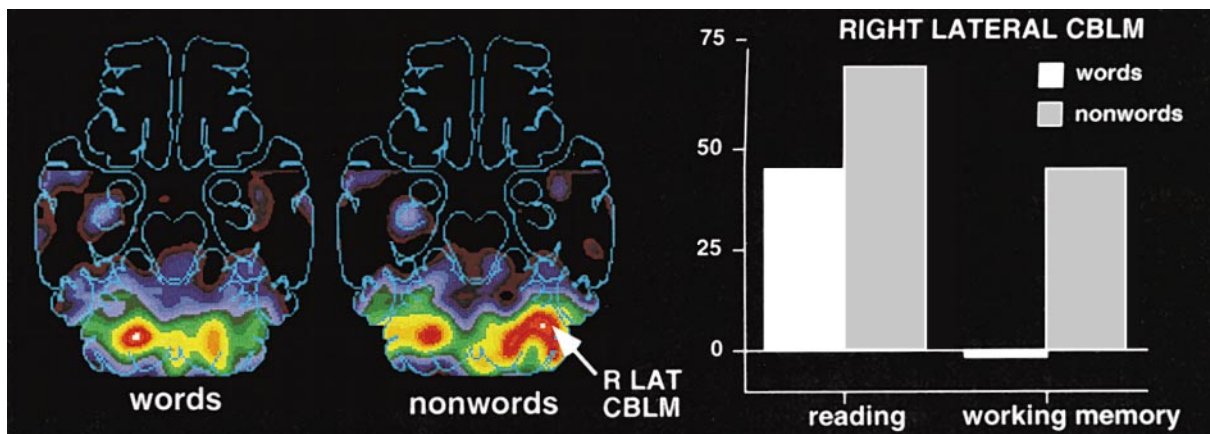

Figure 3. Paramedial and Right Lateral Cerebellar Activation Is Shown for the Overall Word Minus Fixation Condition and the Nonword Minus Fixation Condition

The images are horizontal sections located 16 mm below a plane through the anterior and posterior commissures, with magnitudes represented by a color scale that ranges from 0 (dark purple) to 80 (white) counts. The graphs show mean regional magnitudes across the word and nonword conditions in the present study of reading and in a previously published study of verbal working memory (Fiez et al., 1996).

z = 10). A similarly located focus was found in the overall word minus fixation image (x = −49, y = −33, z = 10), though this region failed to reach significance in the previous analyses. In a direct comparison of the word and nonword conditions, no significant activation difference was found at this region (t[1,10] = 0.87, p = 0.41). The second positive region was located in the right lateral cerebellum, at x = 33, y = −67, z = −14 (Figure 3). Again, a similarly located region could be found in the overall word minus fixation image (x = 33, y = −67, z = −14), but for this region a paired t test revealed greater activation during nonword as compared to word reading (t[1,10] = 2.79; p = 0.03).

Evaluation of Lexicality Effects in a Related Data Set

The differences between words and nonwords were further assessed by evaluating another data set in which subjects were presented with words and nonwords (Fiez et al., 1996). In this previously reported study of verbal working memory, subjects attempted to remember, during the duration of a PET scan, five verbal items presented immediately prior to the beginning of the scan. Each subject performed this task three times, with three different sets of stimulus items: related words, unrelated words, or nonwords. The total data set consisted of 36 short-term memory minus fixation individual difference images (three stimulus conditions from each of 12 subjects).

As regions of interest, we selected those regions that showed a word/nonword difference during reading: a region in the left inferior frontal gyrus (x = −35, y = 15, z = 6), the left fusiform gyrus (x = −43, y = −45, z = −8), and the right lateral cerebellum (x = 33, y = −67, z = −14). Magnitude values within each of these regions were computed for the 36 individual difference images from the working memory study. The values from the related and unrelated word conditions from each subject were averaged together, and then a one-tailed paired t test was used to compare the activation between the word and nonword conditions. For comparison of activation foci, the 12 nonword condition images from the working memory study were averaged together, and in this mean image the nearest focus to each region of interest from the present study was identified.

In the left frontal region, the mean regional activation for nonwords was 32 counts, in contrast to a mean of −2 for words (t[11] = 3.78, p = 0.0015); the nearest focus of activation in the nonword maintenance minus fixation condition was located at x = −29, y = 17, z = 4 (peak magnitude of 57 counts). In the right cerebellar region, the mean region activation for nonwords was 36 counts, in contrast to 13 for words (t[11] = 2.45, p = 0.016); the nearest focus of activation was located at x = 25, y = −67, z = −16 (peak magnitude of 52 counts). In the left fusiform region, no significant difference (t[11] = 0.96, p = 0.18) was found between nonwords and words (means of 13 and 6, respectively). Thus, the lexicality effects (nonwords greater than words) in the left inferior frontal gyrus and right cerebellar regions replicated across studies, but not the effect in the left fusiform region.
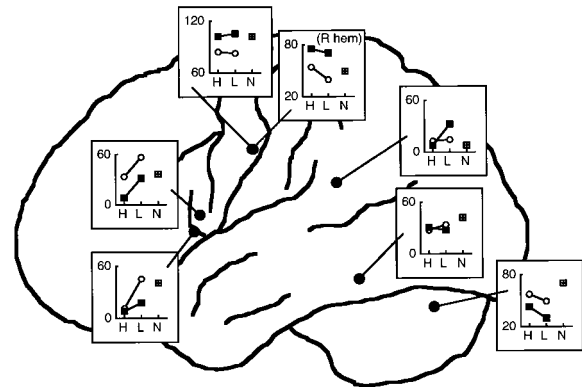

Figure 4. Schematic Diagram Depicting Locations of the Seven Regions that Showed Effects of Frequency, Consistency, or Lexicality

The graphs show mean regional magnitudes across frequency (H, high; L, low; N, nonword), with inconsistent words indicated by the open circles, consistent words indicated by the closed squares, and nonwords indicated by the patterned square.

Discussion

The replication and reliability analyses of the word and nonword conditions identified 18 positive regions of change. These results are consistent with the typical regions found in other neuroimaging studies of reading (reviewed by Fiez and Petersen, 1998). The most relevant finding for the present study is that 7 of the 18 regions showed effects of stimulus condition, ranging from simple main effects of frequency, consistency, or lexicality, to interactions between these factors (see Figure 4). Therefore, even though the brain regions involved in reading may form a distributed and highly interactive network, the network is not so interactive that all regions within the network show highly correlated activity. Specific patterns of regional activation can lead to new hypotheses about the types of cognitive processes supported by different brain regions, and they can be used to constrain theoretical models of reading. To illustrate this point, we will focus our discussion on a region located near the junction of the inferior frontal gyrus and the anterior insula, and bilateral primary motor cortex. These regions were chosen because their patterns of stimulus-related activation are very reliable, they show different relationships to subject performance, and they motivate different theoretical accounts for the contributions of these regions to reading.

Effects of Frequency, Consistency, and Lexicality and Left Inferior Frontal Cortex

The pattern of activation in a left medial inferior frontal region (at or near BA 45, and the anterior insula) paralleled the behavioral data: namely, the region showed an interaction between frequency and consistency, with strong activation only in the low-frequency inconsistent condition. The region also showed an effect of lexicality, with greater activation for nonwords than words. These results are remarkably consistent, both in terms of pattern and location of activation, with three other studies of word reading that have included words and nonwords

Table 7. Left Inferior Frontal Activation across Studies

Study	Low-Frequency Words		Nonwords
	Consistent	Inconsistent	
Fiez et al., this study	not significant	−35, 15, 6	−39, 11, 12
Hagoort et al., 1999	not tested	not tested	−46, 17, −8[a]
Herbster et al., 1997	not significant	−40, 12, −4	−48, 6, 0
Rumsey et al., 1997	not tested	−32, 8, 4	−32, 16, 4[b]

[a] From comparison of nonwords to high-frequency consistent words.
[b] From pseudohomophone condition.

(Herbster et al., 1997; Rumsey et al., 1997; Hagoort et al., 1999) (see Table 7).

Previous authors have suggested that the pattern of activation in this region argues in favor of connectionist models of word reading (Herbster et al., 1997; Rumsey et al., 1997). In connectionist models (Seidenberg and McClelland, 1989; Plaut et al., 1996), two factors drive the development of a distributed phonological representation from an orthographic input. The first is the degree of prior experience with the specific item (providing an influence of frequency), and the second is the degree of prior experience with visually similar words with similar mappings onto phonology (providing an influence of consistency). The settling time and accuracy of the model's responses show a frequency by consistency interaction and an effect of lexicality, which is presumably mirrored in the processing of the constituent layers involved in orthographic to phonological transformation. Since the left frontal region shows a similar sensitivy to frequency, consistency, and lexicality, it may reflect a neural instantiation of this type of distributed processing.

However, we believe that a dual route framework (e.g., Marshall and Newcombe, 1973; Morton and Patterson, 1980; Coltheart et al., 1983; Besner and Smith, 1992; Monsell et al., 1992; Coltheart et al., 1993) could also account for the left frontal activation pattern. The region could be involved in the assembled route, in which a rule-based process is used to "sound out" a pronunciation on the basis of correspondences between individual graphemes and phonemes, or it could be a recipient of information from both the assembled and the direct routes (e.g., Coltheart, 1993). "Activation" of the assembled route in the nonword condition would be necessary because the lexical route cannot support generalizations to nonwords. For consistent words, little activation of the assembled route would be expected, because both routes contribute to a correct response. For low-frequency inconsistent words, this is not the case: the output from the assembled route will actually produce incorrect information. Until the competition between the two routes can be resolved, continued "activation" of the assembled route, or a recipient buffer, might be necessary.

Both types of models contain reading-specific and more general phonological components to which the frontal region might be mapped. Converging data indicate that the region serves as a more general phonological component that contributes to, but is not limited to, orthographic to phonological transformation. Previous neuroimaging studies have identified similar left frontal activation across a range of phonological tasks, some of which did not require orthographic to phonological transformation (e.g., Zatorre et al., 1992; Fiez et al., 1995). In the present study, converging lexicality effects were found in a reading and a verbal working memory task. In the working memory task, orthographic to phonological transformation should have been completed prior to the acquisition of the neuroimaging data. Thus, it is most likely that the lexicality effect arises in the context of a phonologically based rehearsal strategy used to maintain the presented items in working memory (Paulesu et al., 1993; Fiez et al., 1996; Jonides et al., 1998).

Turning to the neuropsychological literature, there are associations between frontal lobe damage and phonological dyslexia, which is characterized by a lexicality effect in word reading (impaired nonword reading in the context of relatively preserved word reading) (reviewed by Fiez and Petersen, 1998). Phonological dyslexia traditionally has been viewed as a highly specific deficit in orthographic to phonological transformation, but recent work indicates that most phonological dyslexics have difficulty on a range of phonological tasks. For instance, those with frontal damage can usually repeat a single nonword immediately after it is spoken, but their performance breaks down when they are asked to respond after a short delay. Deficits in delayed repetition show a lexicality effect, with poorer performance for nonwords than words (Patterson and Marcel, 1992; Farah et al., 1996; Patterson et al., 1996; Sasanuma et al., 1996).

At the same time that the neuropsychological and neuroimaging data converge to implicate the left frontal region in a general form of phonological processing, they also suggest a functional specificity that is not accounted for by either dual route or connectionist frameworks. In both frameworks, the phonological output component is a "bottleneck" that must be accessed to read a word or to do anything else that requires a phonological representation. In contrast, both neuroimaging and neuropsychological data indicate that the left frontal gyrus is not critical for all types of phonological processing. In the present study, little activation of the region was observed in the high-frequency word conditions (Petersen et al., 1989). More importantly, phonological dyslexics with left frontal damage are able to read most words accurately, and word reading is usually intact in Broca's aphasics with left frontal damage (Benson, 1979; Rosen et al., submitted). One possible interpretation of these results is that the left frontal region is specifically involved in the effortful retrieval, manipulation, or selection of a phonological representation, through computational mechanisms that may be similar to those incorporated into either dual route or connectionist models. This interpretation is consistent with neuroimaging evidence that different subregions within the left inferior frontal gyrus are specialized for different types of controlled language processing (Demb et al., 1995; Fiez, 1997) and with behavioral evidence that subjects can use phonological information strategically (McQuade, 1981; Monsell et al., 1992; Balota and Ferraro, 1996).

Effects of Consistency and Primary Motor Cortex
Primary motor cortex showed a pattern of activation that differed from that observed in the left frontal region.

Specifically, effects of consistency, which did not interact with frequency, were found bilaterally, and similar activation was found for low-frequency consistent words and nonwords. These results converge with those reported by Rumsey et al. (1997), who found greater activation for nonwords than low-frequency inconsistent words in primary motor cortex bilaterally. The findings in motor cortex represent an interpretational challenge for several reasons. First, the activation in these areas does not track with reaction time or accuracy, and thus it cannot be explained by a simple mapping to behavior. Second, the common assumption is that frequency and consistency effects arise solely during the process of accessing phonological representations within the language system. Frequency and consistency effects on motor production are not incorporated into either dual route or connectionist models, even though behavioral evidence indicates that such effects exist (Balota and Chumbley, 1985; Inhoff and Topolski, 1994; Seidenberg et al., 1996).

Interpretations of the consistency effect fall into three general categories. One possibility is that primary motor cortex contributes to the computations involved in orthographic to phonological transformation. For instance, word naming may be such an overlearned skill that under some conditions the pronunciation is based largely upon mappings between visual word forms and articulatory programs; semantic and phonological representations may, in large part, be bypassed. From this perspective, reading aloud may involve orthographic to phonological and/or orthographic to articulatory transformation. Inconsistent words may rely more upon a "direct" orthographic to articulatory route, for reasons analogous to those outlined for the direct route in dual route frameworks. Consequently, the motoric programming for inconsistent words may become especially efficient and may produce a smaller increase in regional blood flow. Unlike the left inferior frontal region, motor cortex may not be sensitive to conflict between competing representations, and this could explain the dissociation between frequency and consistency. Computational models could play an invaluable role in understanding dissociations between the activity in different levels of a hierarchical, but interactive, processing stream.

A second possibility is that motor cortex is "downstream" from the regions and processes involved in orthographic to phonological transformation, but factors that influence these regions propagate through to the speech production system. For instance, activity may begin building in motor regions as soon as partial phonological information is available. Since the activation levels in motor regions did not track reaction time, simplistic models of this build-up are unlikely (e.g., activation builds until it passes a threshold). Complex interactions, such as inhibition of regularized pronunciations for inconsistent words, are more likely. Once again, computational models could play an invaluable role in evaluating the plausibility of this account.

A third possibility, which we favor, is that the consistency effects do not reflect an influence of orthography upon phonology, but rather reflect a covarying factor that affects motoric aspects of response initiation or articulation. For instance, if motor cortex represents articulatory gestures at the syllable or phoneme sequence level (Levelt and Wheeldon, 1994), then the effects of consistency may actually represent an effect of frequency at a sublexical level. Interestingly, a recent study involving learned finger sequences provides evidence in support of increased activity in primary motor cortex for highly learned sequences (Karni et al., 1998).

The available data do not permit us to distinguish between the different possibilities. However, it should be noted that all three involve speculations about phonological/articulatory representations and processes that are not incorporated into either dual route or connectionist theoretical frameworks. Furthermore, they can be used to generate a set of testable predictions. For instance, the consistency effects observed in motor cortex could be evaluated through behavioral studies using delayed pronunciation tasks (Inhoff and Topolski, 1994), through neuroimaging work involving auditory presentation and spoken repetition of the same stimulus items and through modeling work exploring the feasibility of maintaining an interaction between frequency and consistency in the model's output, while incorporating consistency-dependent representations of articulation. Thus, the pattern of activation in primary motor cortex exemplifies the ways in which neuroimaging data can draw attention to alternative theoretical accounts of cognitive task performance.

Conclusions

Our results demonstrate that stimulus manipulations can be used to validate, constrain, and stimulate alternative theoretical accounts of the cognitive processes and neural substrates that support reading. To illustrate this point, we considered activation within left inferior frontal and bilateral primary motor cortex. The left inferior frontal region appears to contribute to orthographic to phonological transformation and other phonological tasks, as evidenced in part by the observation of lexicality effects in both a reading and a working memory task. Existing models of reading fail to account for these findings completely because they either localize lexicality effects to a reading-specific component or to a phonological component that is so general it cannot explain the minimal frontal involvement in some tasks (e.g., reading high-frequency words). The pattern of activation in motor cortex showed an effect of consistency, which challenges the common assumption that frequency and regularity effects arise primarily from processing within the language system. Three different hypotheses were developed to account for these consistency effects; all three would require theoretical models of reading to be revised or extended, because they postulate that motor speech areas contribute to stimulus-related differences in reading speed and/or accuracy.

Experimental Procedures

Subjects

Strongly right-handed subjects (Raczkowski et al., 1974) were recruited from the undergraduate and medical campuses of Washington University. They were paid $75 for their participation and gave informed consent according to guidelines set by the Human Studies and Radioactive Drug Research Committees of Washington University. Data were analyzed from 11 subjects: five females and six males, with a mean age of 22 $\pm$ 3 SD years.

Experimental Paradigm

Nine scans were completed in each subject. For three scans (scans 2, 5, and 8), subjects performed a visual fixation control task, in which they maintained fixation on a 3×3 mm black cross hair. For four scans, subjects read aloud one of four word types: high-frequency consistent, low-frequency consistent, high-frequency inconsistent, or low-frequency inconsistent words. For two scans, subjects read aloud pronounceable nonwords. Subjects were given four examples of the types of words to expect before each scan. The reading tasks were performed during scans 1, 3, 4, 6, 7, and 8, counterbalancing the condition order across subjects.

Word and nonword stimuli were presented 3 mm below the fixation cross hair, in uppercase letters subtending a visual angle of approximately 0.6° per letter. Stimuli were displayed for 150 ms, at a rate of one every 1500 ms. An amplitude-gated electronic circuit (voice key) was interfaced with a Macintosh IIci computer system and used to measure voice onset times.

Stimulus Materials

The stimuli were divided into four 50 item word lists (available at http://www.neuron.org/cgi/content/full/24/1/205/DC1): a high-frequency consistent, a high-frequency inconsistent, a low-frequency consistent, and a low-frequency inconsistent word list. The consistent words can be pronounced correctly according to the spelling-to-sound rules of English (Venezky, 1970). The inconsistent words included words for which the spelling-to-sound rules of English produce incorrect pronunciations (e.g., *pint*), and unusual words for which there are few, if any, words with a similar spelling (e.g., *chaos*).

The high-frequency consistent and inconsistent lists had a median frequency of 144 and 145 occurrences per million words (Kucera and Francis, 1967), with respective means of 353 ± 598 SD and 361 ± 496 SD. Both the low-frequency consistent and inconsistent lists had a median frequency of 5, with respective means of 26 ± 136 SD and 8 ± 8 SD (the elevated frequency in the low-frequency inconsistent list reflected the accidental inclusion of *through*, instead of *thorough*; without this item, the mean frequency was 7 ± 8 SD). The lists were also matched for other factors, including mean syllable length (which ranged from 1.2–1.3), and mean letter length (4.7 for all four lists).

Nonwords were created by substituting one or more of the graphemes in each word to create a nonword (e.g., VADE was derived from VASE), while also keeping the distribution of phonemes in the word and matched nonword lists similar. Every subject saw half of the nonword items, divided into two separate lists. For each nonword list seen by a given subject, one quarter of the items were derived from each of the four word lists.

Behavioral Analysis

Word responses that matched the preferred pronunciation were scored as correct (Websters Ninth Collegiate Dictionary; Mish et al., 1996). Incorrect responses were classified according to three different categories: (1) atypical responses for which there is some spelling to sound correspondence, including less preferred pronunciations and regularizations of inconsistent words, (2) other incorrect responses, such as deleting a phoneme or substituting a visually similar word, and (3) dysfluent responses, in which the correct response was preceded by an extraneous utterance.

The most common response to each nonword was determined and scored as correct; in all cases, these were accurate rule-based translations (Venezky, 1970). All other pronunciations were classified using the same error categories developed for the words. Pronunciations that rhymed with a graphemically similar inconsistent word (e.g., pronouncing *stull* such that it rhymes with *pull*) were classified as atypical.

Acquisition and Preprocessing of PET Data

The PETT VI system was employed in the low-resolution mode, which simultaneously acquired seven parallel slices with a center-to-center distance of 14.4 mm (Ter-Pogossian et al., 1982; Yamamoto et al., 1982). For each scan, ^{15}O-labeled water was administered as an intravenous bolus of 8–10 ml of saline containing 48–74 mCi (Welch and Kilbourn, 1985). Data were acquired over a 40 s interval and reconstructed using filtered back-projection to create

images with an in-plane resolution of 17 mm. Because of the linear relationship between tissue radioactivity and blood flow (Herscovitch et al., 1983; Fox and Mintun, 1989), images of radioactive counts and not blood flow were used. An interval between scans (approximately 10 min) allowed radioactivity levels to return to baseline levels.

To control for global fluctuations in activity and variations in amount of isotope injected (Fox et al., 1987), the mean number of counts measured in each image was normalized to 1000. To allow comparisons to be made across subjects, each normalized image was linearly interpolated to form a 49 slice image, which was then cast into the space of the Talairach and Tournoux (Talairach and Tournoux, 1988) stereotactic atlas on the basis of dimensions obtained from a lateral skull X ray (Fox et al., 1985, 1988).

Analysis of Imaging Data

The interpolated and anatomically transformed images from a single scan were used to create individual difference images by using data from a single subject and either (1) taking one of reading task scan images and subtracting a single fixation control scan image or (2) using an image weighting procedure (Shulman et al., 1997) to create a difference image of a single reading task scan minus the weighted average all of the useable fixation control scans from the same subject. The first procedure was used for the hypothesis generation/test analyses in order to create independent data sets, while the second procedure was used for all other analyses. Individual difference images were averaged together to create several different mean difference images; for instance, an overall word reading minus fixation mean difference image was created by averaging the 44 word reading minus fixation individual difference images. Mean difference images were used to identify foci of change during a particular task condition, using an automated peak detection algorithm (Mintun et al., 1989). A relatively low-resolution search that required foci to be separated by more than 16 mm (a criterion reflecting the image resolution) was used to define regions of interest. A higher-resolution search that considered all foci was used to find foci nearest to previously described regions of interest, since there is greater precision involved in localizing versus resolving foci (Fox et al., 1987). Identified foci were used to define the center of spherical regions of interest with 7 mm radii. For each region of interest, individual regional magnitude values were computed by calculating the average magnitude of all pixels located inside the region, in each of a set of individual difference images, excluding data from subjects in which the region was poorly ($<70\%$) sampled. These regional magnitude values were then submitted to standard statistical analysis.

References

Andrews, S. (1982). Phonological recoding: is the regularity effect consistent? Mem. Cognit. *10*, 565–575.

Balota, D.A., and Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. J. Exp. Psychol. Hum. Percept. Perform. *10*, 340–357.

Balota, D.A., and Chumbley, J.I. (1985). The locus of word-frequency effects in the pronunciation task: lexical access and/or production? J. Mem. Lang. *24*, 89–106.

Balota, D.A., and Ferraro, R.F. (1996). Lexical, sublexical, and implicit memory processes in healthy young, and healthy older adults and in individuals with dementia of the Alzheimer type. Neuropsychology *10*, 82–95.

Benson, D.F. (1979). Aphasia, Alexia, and Agraphia (New York: Churchill Livingstone).

Besner, D. (1999). Basic processes in reading: multiple routines in localist and connectionist models. In Converging Methods for Understanding Reading and Dyslexia, P. McMullen and M. Klein, eds. (Cambridge, MA: MIT Press), in press.

Besner, D., and Smith, M.C. (1992). Models of visual word recognition: when obscuring the stimulus yields a clearer view. J. Exp. Psychol. Learn. Mem. Cogn. 18, 468–482.

Buckner, R.L., Raichle, M.E., Miezin, F.M., and Petersen, S.E. (1996). Functional anatomic studies of memory retrieval for auditory words and visual pictures. J. Neurosci. 16, 6219–6235.

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term memory recall of repeated and novel word lists. Mem. Cognit. 21, 539–545.

Coltheart, M., Masterson, J., Byng, S., Prior, M., and Riddoch, J. (1983). Surface dyslexia. Q. J. Exp. Psychol. 35A, 469–495.

Coltheart, M., Curtis, B., Atkins, P., and Haller, M. (1993). Models of reading aloud: dual route and parallel-distributed-processing approaches. Psychol. Rev. 100, 589–608.

Demb, J.B., Desmond, J.E., Wagner, A.D., Vaidya, C.J., Glover, G.H., and Gabrieli, J.D.E. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and and process specificity. J. Neurosci. 15, 5870–5878.

Farah, M.J., Stowe, R.M., and Levinson, K.L. (1996). Phonological dyslexia: loss of a reading-specific component of the cognitive architecture? Cognit. Neuropsychol. 13, 849–868.

Fiez, J.A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. Hum. Brain Map. 5, 79–83.

Fiez, J.A., and Petersen, S.E. (1998). Neuroimaging studies of word reading. Proc. Natl. Acad. Sci. USA 94, 914–921.

Fiez, J.A., Tallal, P., Raichle, M.E., Miezin, F.M., Katz, W.F., and Petersen, S.E. (1995). PET studies of auditory and phonological processing: effects of stimulus characteristics and task demands. J. Cogn. Neurosci. 7, 357–375.

Fiez, J., Raife, E., Balota, D., Schwarz, J., Raichle, M., and Petersen, S. (1996). A positron emission tomography study of the short-term maintenance of verbal information. J. Neurosci. 16, 808–822.

Forster, K., and Chambers, S. (1973). Lexical access and naming time. J. Verb. Learn. Verb. Behav. 12, 627–635.

Fox, P.T., and Mintun, M.A. (1989). Noninvasive functional brain mapping by change-distribution analysis of averaged PET images of $H_2^{15}O$. J. Nucl. Med. 30, 141–149.

Fox, P.T., Perlmutter, J.S., and Raichle, M.E. (1985). A stereotactic method of anatomical localization for positron emission tomography. J. Comput. Assist. Tomogr. 9, 141–153.

Fox, P.T., Miezin, F.M., Allman, J.M., Van Essen, D.C., and Raichle, M.E. (1987). Retinotopic organization of human visual cortex mapped with positron emission tomography. J. Neurosci. 7, 913–922.

Fox, P.T., Mintun, M.A., Reiman, E.M., and Raichle, M.E. (1988). Enhanced detection of focal brain responses using intersubject averaging and change-distribution analysis of subtracted PET images. J. Cereb. Blood Flow Metab. 8, 642–653.

Frederiksen, J., and Kroll, J. (1976). Spelling and sound: approaches to the internal lexicon. J. Exp. Psychol. Hum. Percept. Perform. 2, 361–379.

Glushko, R.J. (1979). The organization and activation of orthographic knowledge in reading aloud. J. Exp. Psychol. Hum. Percept. Perform. 5, 674–691.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., and Seitz, R.J. (1999). The neural circuitry involved in the reading of German words and pseduowords: a PET study. J. Cogn. Neurosci., in press.

Herbster, A.N., Mintun, M.A., Nebes, R.D., and Becker, J.T. (1997). Regional cerebral blood-flow during word and non-word reading. Hum. Brain Map. 5, 84–92.

Herscovitch, P., Markham, J., and Raichle, M.E. (1983). Brain blood flow measured with intravenous H215O. I. Theory and error analysis. J. Nucl. Med. 24, 782–789.

Hunton, D.L., Miezin, F.M., Buckner, R.L., van Mier, H.I., Raichle, M.E., and Petersen, S.E. (1996). An assessment of functional-anatomical variability in neuroimaging studies. Hum. Brain Map. 4, 122–139.

Inhoff, A.W., and Topolski, R. (1994). Use of phonological codes during eye fixation in reading an in on-line and delayed naming tasks. J. Mem. Lang. 33, 689–713.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. J. Mem. Lang. 36, 505–529.

Jescheniak, J.D., and Levelt, W.J.M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. J. Exp. Psychol. Learn. Mem. Cogn. 20, 824–843.

Jonides, J., Schumacher, E.H., Smith, E.E., Koeppe, R.A., Awh, E., Reuter-Lorenz, P.A., Marshuetz, C., and Willis, C.R. (1998). The role of parietal cortex in verbal working memory. J. Neurosci. 18, 5026–5034.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M.M., Turner, T., and Ungerleider, L.G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proc. Natl. Acad. Sci. USA 95, 861–868.

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

Levelt, W.J., and Wheeldon, L. (1994). Do speakers have access to a mental syllabary? Cognition 50, 239–269.

Marcel, A.J. (1980). Surface dyslexica and beginning reading: a revised hypothesis of the pronunciation of print and its impairments. In Deep Dyslexia, M. Coltheart, K. Patterson, and J.C. Marshall, eds. (London: Routledge and Kegan Paul), pp. 227–258.

Marshall, J.C., and Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. J. Psycholing Res. 2, 175–198.

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

Meyer, D.E., Schvaneveldt, R.W., and Ruddy, M.G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. Mem. Cognit. 2, 309–321.

Mintun, M.A., Fox, P.T., and Raichle, M.E. (1989). A highly accurate method of localizing regions of neuronal activation in the human brain with positron emission tomography. J. Cereb. Blood Flow Metab. 9, 96–103.

Mish, F., Gilman, E.W., Lowe, J.G., McHenery, R.D., and Pease, R.W., eds. (1996). Webster's Ninth New Collegiate Dictionary. (Springfield, MA: Merriam-Webster).

Monsell, S., Patterson, K., Graham, A., Hughes, C.H., and Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: strategic anticipation of lexical status. J. Exp. Psychol. Learn. Mem. Cogn. 18, 452–467.

Morton, J., and Patterson, K. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In Deep Dyslexia, M. Coltheart, K. Patterson, and J.C. Marshall, eds. (London: T.J. Press), pp. 91–118.

Patterson, K., and Marcel, A. (1992). Phonological ALEXIA or PHONOLOGICAL alexia? In Analytic Approaches to Human Cognition, J. Alegria, D. Holdender, J.J. de Morais, and M. Radeua, eds. (Amsterdam: Elsevier Science Publishers), pp. 259–274.

Patterson, K., Suzuki, T., and Wydell, T. (1996). Interpreting a case of Japanese phonological alexia: the key is in phonology. Cognit. Neuropsychol. 13, 803–822.

Paulesu, E., Frith, C.D., and Frackowiak, R.S.J. (1993). The neural correlates of the verbal component of working memory. Nature 362, 342–345.

Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., and Raichle, M.E. (1989). Positron emission tomographic studies of the processing of single words. J. Cogn. Neurosci. 1, 153–170.

Plaut, D.C., McClelland, J.F., Seidenberg, M.S., and Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. Psychol. Rev. 103, 56–115.

Pugh, K.R., Shaywitz, B.A., Shaywitz, S.E., Shankweiler, D.P., Katch,

L., Fletcher, J.M., Skudlarski, P., Fulbright, R.K., Constable, R.T., Bronen, R.A., et al. (1997). Predicting reading performance from neuroimaging profiles: the cerebral basis of phonological effects in printed word identification. J. Exp. Psychol. Hum. Percept. Perform. *23*, 299–318.

Raczkowski, D., Kalat, J.W., and Nebes, R. (1974). Reliability and validity of some handedness questionnaire items. Neuropsychologia *12*, 43–47.

Rumsey, J.M., Horwitz, B., Donohue, C., Nace, K., Maisog, J.M., and Andreason, P. (1997). Phonologic and orthographic components of word recogntion: a PET-rCBF study. Brain *119*, 739–759.

Sasanuma, S., Ito, H., Patterson, K., and Ito, T. (1996). Phonological alexia in Japanese: a case study. Cognit. Neuropsychol. *13*, 823–848.

Scarborough, D.L., Cortese, C., and Scarborough, H.S. (1977). Frequency and repetition effects in lexical memory. J. Exp. Psychol. Hum. Percept. Perform. *3*, 1–17.

Seidenberg, M.S., and McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. Psychol. Rev. *96*, 523–568.

Seidenberg, M.S., Petersen, A., and MacDonald, M.C. (1996). Pseudohomophone effects and models of word recognition. J. Exp. Psychol. Learn. Mem. Cogn. *22*, 48–62.

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

Shulman, G.L., Corbetta, M., Buckner, R.L., Raichle, M.E., Fiez, J.A., Miezin, F.M., and Petersen, S.E. (1997). Top-down modulation of early sensory cortex. Cereb. Cortex *7*, 1047–3211.

Taft, M., and van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. J. Mem. Lang. *38*, 203–224.

Talairach, J., and Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain. M. Rayport, trans. (New York: Thieme Medical Publishers).

Taraban, R., and McClelland, J.L. (1987). Conspiracy effects in word pronunciation. J. Mem. Lang. *26*, 608–631.

Ter-Pogossian, M.M., Ficke, D.C., Hood, J.T., Yamamoto, M., and Mullani, N.A. (1982). PETT VI: a positron emission tomograph utilizing cesium fluoride scintillation detectors. J. Comput. Assist. Tomogr. *6*, 125–133.

Venezky, R.L. (1970). The Structure of English Orthography (The Hague, the Netherlands: Mouton).

Welch, M., and Kilbourn, M. (1985). A remote system for the routine production of oxygen-15 radiopharmaceuticals. J. Labeled Comp. Radiopharm. *22*, 1193–1200.

Yamamoto, M., Ficke, D.C., and Ter-Pogossian, M. (1982). Performance study of PETT VI, a positron computed tomograph with 288 cesium fluoride detectors. IEEE Trans. Nucl. Sci. *NS-29*, 529–533.

Zatorre, R.J., Evans, A.C., Meyer, E., and Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. Science *256*, 846–849.