fMRI Evidence for Dual Routes to the Mental Lexicon in Visual Word Recognition

Christian J. Fiebach, Angela D. Friederici, Karsten Müller, and D. Yves von Cramon

Abstract

Event-related fMRI was used to investigate lexical decisions to words of high and low frequency of occurrence and to pseudowords. The results obtained strongly support dual-route models of visual word processing. By contrasting words with pseudowords, bilateral occipito-temporal brain areas and posterior left middle temporal gyrus (MTG) were identified as contributing to the successful mapping of orthographic percepts onto visual word form representations. Low-frequency words and pseudowords elicited greater activations than high-frequency words in the superior pars opercularis [Brodmann’s area (BA) 44] of the left inferior frontal gyrus (IFG), in the anterior insula, and in the thalamus and caudate nucleus. As processing of these stimuli during lexical search is known to rely on phonological information, it is concluded that these brain regions are involved in grapheme-to-phoneme conversion. Activation in the pars triangularis (BA 45) of the left IFG was observed only for low-frequency words. It is proposed that this region is involved in processes of lexical selection.

INTRODUCTION

In the present study, we provide neuroimaging evidence concerning the different roles that occipito-temporal and left inferior frontal brain areas play during visual word processing in a task requiring nothing more than a brief access to the mental lexicon, namely the lexical decision task. Lexical decisions to pseudowords are known to proceed more slowly than those to words, reflecting the need for a more exhaustive search in the mental lexicon when no lexical entry is found (e.g., Forster & Bednall, 1976). Low-frequency words, as compared to high-frequency words, also elicit increased decision times (e.g., Allen, Wallace, & Weber, 1995). The word frequency effect in lexical decisions has been attributed to a prolonged serial search in a frequency-ordered lexicon in so-called serial search models (e.g., Forster, 1976). Alternatively, it has been explained by differences in the weights of distributed connections within the network representing the mental lexicon in connectionist network models (e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989). Within the latter framework, the strength of the connections is tuned by the frequency of exposure to an item.

On the basis of such models, one might posit that brain regions involved in lexical access can be identified by measuring neural activity elicited during the processing of words of different frequencies. So far, there is only one neuroimaging study in which word frequency, among other variables, was systematically manipulated (Fiez, Balota, Raichle, & Petersen, 1999). This PET study used a word-naming paradigm. Effects of word frequency were found in the left superior temporal gyrus [Brodmann’s area (BA 22) and in the supplementary motor area (SMA/BA 6)]. Furthermore, the left inferior frontal gyrus (IFG; BAs 44/45), extending to the anterior insula, showed an interaction of regularity with frequency due to increased activity for low-frequency words with inconsistent spelling-to-sound mapping. The analogy of this finding to the empirically well-established interaction of regularity and frequency in naming times (e.g., Seidenberg, Waters, Barnes, & Tanenhaus, 1984) was interpreted by Fiez et al. (1999) as suggesting a contribution of left inferior frontal cortex to rule-based phonological processes of sounding out low-frequency words with inconsistent spelling-to-sound mapping. In general, the left inferior frontal region was assumed to be involved in effortful processes such as retrieval, manipulation, or selection of phonological representations (Fiez et al., 1999).

According to dual-route models (e.g., Coltheart, Curtis, Atkins, & Haller, 1993), the functional mechanisms of word reading involve a fast pathway to the mental lexicon that directly maps orthographic percepts onto stored word form representations. In parallel, a slower “assembled” or nonlexical route translates the graphemic input into phonological information. The grapheme-
to-phoneme conversion route has been postulated in order to explain the ability to read aloud pronounceable nonwords, which do not have a stored phonological representation. In the study of Fiez et al. (1999), left fronto-insular activations were associated with the grapheme-to-phoneme conversion route.

It has been argued that both graphemic and phonological information are involved in word identification (e.g., Perfetti, 1999). The exact timing of the different contributions of these sources of information depends on a number of variables. In word recognition, both routes will be activated in parallel, and the faster direct route generally wins the race (Paap & Noel, 1991). When comparing words of high and low frequency, lexical search based on visual word form (i.e., the direct route) more rapidly identifies the lexical entries for high-frequency words than for low-frequency words. For low-frequency words, grapheme-to-phoneme conversion rules can compute the phonological form before the lexical entry is identified through the direct route. Grapheme-to-phoneme conversion thus supports lexical search (cf. Perfetti, 1999). In contrast, when no lexical entry is found, as is the case for pseudowords, it is assumed that an exhaustive search of the lexicon is carried out before the decision is made (e.g., Forster & Bednall, 1976). Like in low-frequency words, the phonological form is assembled by grapheme-to-phoneme rules, because the direct route does not yield a fast result. However, in the case of pseudowords, no lexical entry is found in the mental lexicon. According to the logic outlined here, it should be possible to identify the neural substrates of the dual routes to the mental lexicon by directly contrasting neural activity elicited during the processing of pseudowords and of words of high and low frequency in the lexical decision task.

The naming paradigm employed by Fiez et al. (1999) and in many other studies of word processing involves the activation of phonological codes and the articulation of word and nonword stimuli. In the lexical decision task, however, these processes are not necessarily required. Rather than investigating the complete cascade of processes from visual input to articulation, the present study intends to yield insights into the initial stages of visual word processing, including the computation of a visual word form from the orthographic input (e.g., Bub & Arguin, 1995) and the selection of the correct lexical entry. At this point, a lexical decision can be made (e.g., Coltheart, Davelaar, Jonasson, & Besner, 1977; Coltheart et al., 1993). Brain activation observed in the lexical decision paradigm, thus, does not reflect higher-level phonological or semantic processes usually involved in tasks like word naming, rhyme judgments, or semantic judgments. If activation of cortical areas implicated in these higher-order processes is observed in the lexical decision paradigm, it can be concluded that these areas are also coactivated during earlier stages of word processing and thus contribute to lexical access.

In order to investigate how the brain achieves lexical access during visual word processing, we utilized the lexical decision task in an event-related functional magnetic resonance imaging (fMRI) experiment. In the lexical decision task, participants were presented with a pseudorandomly ordered sequence of high- and low-frequency words (i.e., 68 inanimate nouns of each type) and with the same number of phonologically legal pseudowords that were derived from the word stimuli by randomly exchanging one or two letters. Stimuli were presented for 400 msec, separated by an interstimulus interval (ISI) of 7.1 sec. The participants’ task was to judge whether or not each stimulus was a legal German word. During the ISI, a visual feedback was given to the participants. Only correctly answered trials were used for statistical analysis of brain activation data.

The event-related approach allowed us to track the hemodynamic response elicited by the processing of individual items (e.g., Buckner, 1998). When contrasting the summed responses to all trials of the different experimental conditions, task components common to all conditions (such as visual coding of the stimuli or the actual decision component of the task) cancel out and only condition-specific activity remains for interpretation. Strategic effects on the response behavior in different conditions, which might confound activation results in blocked designs of the lexical decision task, are thus avoided. Reported are direct contrasts (a) between the activations elicited by words and by pseudowords, as well as (b) between words of low and high frequency and (c) between pseudowords and low- and high-frequency words, respectively. We expected that the direct comparison of words with pseudowords should result in activations associated with the mapping of orthographic representations of perceived words onto stored word form representations. This mechanism should be active for words of high frequency, as well as for those with low frequency. The comparison of low- and high-frequency words was expected to isolate neural activity associated with the activation of phonological information by grapheme-to-phoneme conversion rules for low-frequency words. Brain regions involved in grapheme-to-phoneme conversion should also be observable for pseudowords when contrasted with high- but not with low-frequency words, because pseudowords are processed more like low-frequency words.

RESULTS

Lexical Decision Performance

Behavioral data (see Table 1) demonstrate that lexical decision times to words were, on average, 127.26 msec faster than those to pseudowords \[F(1,11) = 58.4; p < .0001\]. The interaction between stimulus type
Brain Activation: Lexicality

As Figure 1 shows, increased brain activations for words as opposed to pseudowords were observed bilaterally in occipito-temporal brain regions that overlap with the “ventral” pathway for visual object recognition (“what-pathway”; e.g., Ungerleider & Mishkin, 1982). In the occipital lobe, activation differences were observed in the inferior occipital gyri and the middle occipital gyri bilaterally. These activations extended into the lateral occipital sulci (cf. Table 2). On the basal surface of the brain, the most posterior activations were observed at the border between the occipital gyri and the posterior fusiform gyri in both hemispheres. Along the banks of the collateral sulcus, substantial portions of the fusiform gyri were activated as displayed in Figure 1. The center of gravity of fusiform activity in the right hemisphere was located more laterally, that is, more towards the lateral occipito-temporal sulcus than in the left hemisphere. Basal activation differences terminated in the anterior portion of the collateral sulcus (cf. Figure 1 and Table 2).

Circumscribed significant activation increases for words were also present within the posterior division of the left middle temporal gyrus (MTG; immediately below the posterior superior temporal sulcus) and along the descending leg of the right intraparietal sulcus (cf. Table 2). No activations were obtained that were stronger for pseudowords than for words.

Brain Activation: Word Frequency

Words of low frequency caused stronger activations than high-frequency words in the left IFG (superior portion of pars opercularis/BA 44 and pars triangularis/BA 45; cf. Figure 2 and Table 3) and in subcortical structures such as the (supero-lateral) head of the caudate nucleus and the (anterior) thalamus bilaterally. In addition, activation increases were observed in the anterior insula bilaterally with maxima in the insular pole. In the left hemisphere, insular activations extended laterally into the deep frontal operculum. No reliable activations, which were stronger for high- than for low-frequency words, were observed. Note that there was no systematic relationship between the onset latency of the visual feedback (which varied up to 70 msec between conditions) and any of the observed activation effects.

In order to statistically determine whether low- and high-frequency words differed in the amount of activation elicited in the occipital and temporal areas, which exhibited stronger activation for words than for pseudowords (see the Brain Activation: Lexicality section), a region-of-interest (ROI) analysis was performed. Spherical ROIs (radius 3 mm) were placed at the local maxima of the foci observed for words versus pseudowords (cf. Table 2). Mean ROI-wise contrasts for low-frequency words and for high-frequency words, as contrasted with the corresponding pseudowords, were extracted from individual contrast images (Bosch, 2000). Regionwise planned comparisons revealed that occipito-temporal activations did not differ between words of high and low frequency (most \( F \)'s < 1.7). Three ROIs had \( F \) values greater than 2 (i.e., 2.83, 2.61, and 2.68). However, these comparisons also were not significant (all \( p \)'s > .12).

In addition to the results reported above, pseudowords were found to elicit greater neural activity than high-frequency words but not than low-frequency words in the superior portion of left pars opercularis (i.e., BA 44) and in the anterior left insula (cf. Figure 3 and Table 3). This pattern was very similar to the one obtained for...
low-frequency words relative to high-frequency words. In addition to differences in the left fronto-insular region, pseudowords also elicited increased neural activity in bilateral subcortical regions involving the head of the caudate nucleus and the (anterior) thalamus. Furthermore, a reliable activation in the anterior portion of the right cingulate gyrus (BAs 24/32) was observed (cf. Table 3). No comparable activation effects were
Table 2. Regions of Reliably Stronger Activation for Words than for Pseudowords

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>mm³</th>
<th>Zmax</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Collateral sulcus, anterior portion</strong></td>
<td>36</td>
<td>L</td>
<td>−28</td>
<td>−34</td>
<td>−12</td>
<td>*</td>
<td>3.97</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>R</td>
<td>33</td>
<td>−36</td>
<td>−8</td>
<td>*</td>
<td>3.59</td>
</tr>
<tr>
<td><strong>Fusiform gyrus, anterior portion (T4)</strong></td>
<td>37</td>
<td>L</td>
<td>−32</td>
<td>−45</td>
<td>−7</td>
<td>*</td>
<td>4.10</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>R</td>
<td>35</td>
<td>−48</td>
<td>−10</td>
<td>*</td>
<td>3.77</td>
</tr>
<tr>
<td><strong>Posterior fusiform gyrus (T4)/inferior occipital gyrus (O3)</strong></td>
<td>37</td>
<td>L</td>
<td>−29</td>
<td>−89</td>
<td>1</td>
<td>*</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>R</td>
<td>32</td>
<td>−82</td>
<td>0</td>
<td>*</td>
<td>3.84</td>
</tr>
<tr>
<td><strong>Middle occipital gyrus (O2)/intraoccipital sulcus</strong></td>
<td>18/19</td>
<td>L</td>
<td>−24</td>
<td>−90</td>
<td>13</td>
<td>*</td>
<td>4.18</td>
</tr>
<tr>
<td></td>
<td>18/19</td>
<td>R</td>
<td>31</td>
<td>−81</td>
<td>19</td>
<td>*</td>
<td>4.13</td>
</tr>
<tr>
<td><strong>Occipito-polar cap/inferior occipital gyrus (O3)</strong></td>
<td>18</td>
<td>L</td>
<td>−18</td>
<td>−98</td>
<td>9</td>
<td>*</td>
<td>4.41</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>R</td>
<td>22</td>
<td>−95</td>
<td>2</td>
<td>*</td>
<td>4.27</td>
</tr>
<tr>
<td><strong>Lateral occipital sulcus</strong></td>
<td>19</td>
<td>L</td>
<td>−37</td>
<td>−68</td>
<td>9</td>
<td>*</td>
<td>4.05</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>R</td>
<td>38</td>
<td>−64</td>
<td>10</td>
<td>*</td>
<td>3.81</td>
</tr>
<tr>
<td><strong>Posterior MTG (T2)</strong></td>
<td>21</td>
<td>L</td>
<td>−52</td>
<td>−52</td>
<td>11</td>
<td>238</td>
<td>3.98</td>
</tr>
<tr>
<td><strong>Intraparietal sulcus (descending leg)</strong></td>
<td>19/7</td>
<td>R</td>
<td>26</td>
<td>−61</td>
<td>34</td>
<td>903</td>
<td>4.35</td>
</tr>
</tbody>
</table>

Listed are BAs, Talairach coordinates (cf. Talairach & Tournoux, 1988), cluster sizes, and Z values (uncorrected). For occipital and basal temporal areas, local maxima within extensive activated regions are reported. Overall sizes of the activated clusters were *3,143 mm³, 4,777 mm³, and *4,300 mm³. Cortical gyri are labeled according to Duvernoy (1999).

Figure 2. Effects of word frequency on neural activity during the lexical decision task. Regions showing stronger activations for low-frequency words as opposed to high-frequency words were left pars triangularis of the inferior frontal gyrus (IFG/tr; BA 45), left superior pars opercularis (IFG/op; BA 44), and the anterior insula (alns) bilaterally. Furthermore, the (supero-lateral) head of the caudate nucleus (CN) and the anterior thalamus (Thal) of both hemispheres were stronger activated when low-frequency words were processed. Talairach coordinates, cluster sizes, and Z values are given in Table 3. Sagittal sections displayed in the upper row are located at x = −48 (left) and x = 45 (right), respectively. Sagittal sections in the lower row represent x = −33 and x = 29. Smoothness estimation for the x, y, and z axes: 8.3, 8.5, and 7.8 FWHM; search volume 1,502.1 resels.
obtained in the contrast of pseudowords with low-frequency words.

**DISCUSSION**

Performance data (i.e., decision times and error rates) obtained during the fMRI experiment demonstrated that the behavioral effects of lexicality and word frequency could be replicated in the MR scanner. In the hemodynamic response, the effect of lexicality was evident in stronger bilateral activation for words as compared to pseudowords in occipito-temporal brain areas and in the posterior portion of the left MTG. An effect of word frequency was observed in the brain’s response in left fronto-insular, as well as in subcortical, structures. Low-frequency words and pseudowords elicited greater activity than high-frequency words in the left IFG, in the anterior insula, and bilaterally in the head of the caudate nucleus and the anterior thalamus.

### Occipito-Temporal Brain Regions

Results of the present study are in line with reports of activation differences in the fusiform gyrus between words and pronounceable nonwords in reading aloud (e.g., Herbster, Mintun, Nebes, & Becker, 1997). However, fusiform activations obtained in the present experiment were located more posteriorly. Other studies failed to find reliable activation differences between words and pronounceable nonwords in these areas (e.g., Tagamets, Novick, Chalmers, & Friedman, 2000). A recent PET study of word naming reported the reverse result, namely, stronger activation for pseudowords than for words was obtained in both fusiform gyri (Hagoort et al., 1999). There are two important differences between the present study and previous work. First, event-related fMRI was used, allowing to model hemodynamic responses elicited by individual events. Brain activations in previous studies of word processing, in which blocked designs were used, might have been influenced by strategic effects on task performance or, in some cases, stereotypic response styles. This holds especially for previous studies investigating performance in lexical decision tasks (e.g., Hart, Kraut, Kremen, Soher, & Gordon, 2000; Kiehl et al., 1999; Table 3. Brain Regions Exhibiting Reliably Stronger Activation for Low-Frequency Words than for High-Frequency Words, as well as for Pseudowords than for High-Frequency Words

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>mm$^3$</th>
<th>Z$_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low- versus high-frequency words (frequency)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFG (F3), pars triangularis</td>
<td>45</td>
<td>L</td>
<td>−51</td>
<td>31</td>
<td>10</td>
<td>835</td>
<td>4.17</td>
</tr>
<tr>
<td>IFG (F3), pars opercularis (superior portion)</td>
<td>44</td>
<td>L</td>
<td>−42</td>
<td>1</td>
<td>17</td>
<td>813</td>
<td>3.98</td>
</tr>
<tr>
<td>Anterior insula</td>
<td></td>
<td>L</td>
<td>−29</td>
<td>28</td>
<td>−4</td>
<td>3,398</td>
<td>4.83</td>
</tr>
<tr>
<td>Caudate nucleus (head) and (anterior) thalamus</td>
<td></td>
<td>L/R</td>
<td>−10</td>
<td>−10</td>
<td>7</td>
<td>1,656</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L</td>
<td>−10</td>
<td>−5</td>
<td>14</td>
<td>1,695</td>
<td>4.81</td>
</tr>
<tr>
<td><strong>Pseudowords versus high-frequency words</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFG (F3), pars opercularis (superior portion)</td>
<td>44</td>
<td>L</td>
<td>−47</td>
<td>10</td>
<td>13</td>
<td>636</td>
<td>4.5</td>
</tr>
<tr>
<td>Anterior insula</td>
<td></td>
<td>L</td>
<td>−37</td>
<td>16</td>
<td>0</td>
<td>1,270</td>
<td>4.15</td>
</tr>
<tr>
<td>Anterior cingulate gyrus</td>
<td>24/32</td>
<td>R</td>
<td>7</td>
<td>31</td>
<td>10</td>
<td>946</td>
<td>4.18</td>
</tr>
<tr>
<td>Caudate nucleus (head) and (anterior) thalamus</td>
<td></td>
<td>L/R</td>
<td>−3</td>
<td>−3</td>
<td>12</td>
<td>3,215</td>
<td>4.36</td>
</tr>
</tbody>
</table>

Listed are BAs, Talairach coordinates of maximal activation differences (cf. Talairach & Tournoux, 1988), cluster sizes, and Z values (uncorrected). Cortical gyri are labeled according to Duvernoy (1999).
Madden et al., 1996; Nenov, Halgren, Mandelkern, & Smith, 1994) in which the processing of words and pseudowords could not be dissociated in the way it was done in the present work. In our study, the observed activation differences can be unequivocally attributed to the psycholinguistic manipulations of the stimulus material. Second, differences in the tasks investigated might be partly responsible for different patterns of activation as compared to other studies of word processing. When participants have to name the stimuli (e.g., Fiez et al., 1999; Hagoort et al., 1999), observed activation results are also influenced by processes associated with the actual articulation of the stimuli.

An involvement of basal temporal cortices in language processing has been described previously. The middle portion of the fusiform gyrus has been referred to as “basal temporal language area” (e.g., Bookheimer, Zeffiro, Blaxton, Gaillard, & Theodore, 1995; Lüders et al., 1991), “multimodal language region” (Büchel, Price, & Friston, 1998), or as “visual word form area” in which a perceptually invariant higher-order orthographic unit (i.e., the visual word form) is computed from the visual input (Cohen et al., 2000). Talairach coordinates of the fusiform activation foci reported by Büchel et al. (1998) and by Cohen et al. (2000) deviated less than 1 cm in any direction from the local maximum that we observed in the anterior portion of the left fusiform gyrus. Brain activation in the fusiform gyrus has also been reported to be influenced by stimulus duration and stimulus presentation rate (cf. Price, Moore, & Frackowiak, 1996).

Intracranial recordings of event-related brain potentials (ERPs) from the inferior temporal lobe in humans (Nobre, Allison, & McCarthy, 1994) have demonstrated that at about 200 msec after stimulus onset, ERPs acquired from the posterior fusiform gyrus discriminate between orthographic stimuli (i.e., letter strings) and other visually presented objects (see also Puce, Allison, Asgari, Gore, & McCarthy, 1996, for PET evidence). In a more anterior portion of the inferior temporal lobe, intracranial ERPs were sensitive to the lexicality of the stimuli at about 400 msec post-stimulus onset, leading to the assumption of a functionally specialized word recognition stream within the “ventral” visual pathway. A recent magnetoencephalographic study provided converging evidence by demonstrating activation in left inferior occipito-temporal cortex during word processing especially between 150 and 250 msec after word onset (Kuriki, Takeuchi, & Hirata, 1998). In this study, it was concluded that these brain areas subserve functions prior to lexical–semantic processes, and it was suggested that visual word form representations are stored here. The assumption that the derivation of a visual word form from the sensory input is a prelexical process is in line with studies of neurological patients suffering from the syndrome of pure alexia. It has been proposed (e.g., Bub & Arguin, 1995; Shallice & Saffran, 1986) that in these patients, the synthesis of letters into higher-order perceptual units (i.e., into visual word form representations) is impaired. Despite these deficits in overt identification, subthreshold activation of word representations is gained, which allows to judge the lexicality of presented stimuli (e.g., Bub & Arguin, 1995).

Our data confirm a sensitivity for words within the “ventral” visual pathway. Moreover, the brain regions identified here correspond to the ones identified by intracranial ERP recordings. The temporal resolution of brain imaging data does not allow to determine whether the observed differences in the neural activity elicited by word and pseudoword processing in occipital areas were due to initial processing stages after visual feature extraction or whether the activation differences were due to reciprocal modulation of occipital areas taking place later during processing. However, there is evidence from ERP data that the brain reacts differently to words and pseudowords as early as 112 msec after stimulus onset (Sereno, Rayner, & Posner, 1998). Furthermore, independence of occipito-temporal activation of word frequency suggests that these activation differences were elicited by prelexical processes.

The fusiform activation reported here was also in the direct vicinity of the distributed neural system for face processing (cf. Haxby, Hoffman, & Gobbini, 2000 for an overview). In a model of the neuroanatomical organisation of the face-processing system, these authors proposed that perceptual processing of facial features takes place within the inferior occipital gyri, while the fusiform gyrus was associated with the representation of invariant aspects of faces that contribute to the recognition of face identity (Haxby et al., 2000). This distinction is likely to hold for the processing of visual words as well. Previous work has demonstrated that within occipito-temporal regions, perceptual features of objects and faces (e.g., Puce et al., 1996) and of different categories of objects (e.g., Chao, Haxby, & Martin, 1999) are represented in distinguishable cortical areas. The present results, however, do not allow to determine whether the distinct functional loci within occipital and temporal cortex described in previous studies contribute differently to word recognition. It has been proposed that basal occipital and temporal cortices might represent feature-based semantic information about object form (e.g., Martin & Chao, 2001). There are a number of arguments that make an interpretation of the present activation results with respect to feature-based semantics unlikely. First, about half of the word stimuli used in the present experiment were abstract nouns, making it unlikely that feature-based object semantics were available for all stimuli. Second, this interpretation of the present results would imply that the activations in question were associated with postlexical processes, which, in turn, are assumed to be influenced by word frequency. This, however, was not the case for the occipito-temporal activations observed in the present experiment.
As an alternative, it has been proposed recently that activation in occipito-temporal brain areas reflects expertise with the processing of highly familiar object categories, rather than the access to anatomically distinct category-specific recognition modules (e.g., Gauthier, Skudlarski, Gore, & Anderson, 2000). On the basis of the present findings, this assumption can be extended to written words. Words clearly represent a category of visual objects for which everybody who knows how to read has high expertise. This view converges with the notion that lexical decisions, to a large degree, reflect stimulus familiarity (Balota & Chumbley, 1984). Stimulus familiarity might very well play an important role in the initial stages of visual word processing, especially if the task requires to determine the lexicality of the processed stimuli. On the other hand, this conclusion is qualified by the finding that occipito-temporal activations were independent of word frequency, a variable that is closely related to familiarity.

Our results suggest that when words are processed, neural mechanisms closely related to visual object recognition are activated. In analogy to other domains of visual processing (like face perception; cf. Hasby et al., 2000), it can be hypothesized that occipito-temporal brain regions are subdivided into functionally distinct regions along the posterior–anterior axis. Inferior occipital cortices, known to be involved in perceptually based processes, thus might contribute to the early perception of orthographic units. Fusiform activations, on the other hand, are involved in recognizing the perceptually invariant visual word form and, thereby, provide the visual access code for lexical representations entailing grammatical, semantic, and phonological information (e.g., Cohen et al., 2000). The most anteriorly located inferior temporal activation, along the anterior portion of the collateral sulcus, might subserve functions of semantic processing. As words elicited differential activity from pseudowords in all these basal areas, the present data cannot dissociate the different roles that these areas play during visual word processing. It remains to be determined whether this suggested functional organization of the basal temporal and occipital cortex, in the context of word processing, can be empirically supported.

The fact that activation in the fusiform gyri was not influenced by the manipulation of word frequency supports the assumption that these regions contribute to an intermediate stage of word processing, namely, before lexical information becomes relevant. Cohen et al. (2000) assumed that within the ventral visual pathway, visual word forms are computed by assembling letters into higher-order units. However, if a visual-orthographic unit would be computed from the input string, this process should elicit comparable activation for lexical and nonlexical stimuli. Instead, there were clear activation differences observed between these two categories in the present study. It appears to be more likely that the fusiform activation reflects the recognition of visual word forms. This interpretation of our findings is in line with the dual-route theory (e.g., Coltheart et al., 1993). In this model, it is assumed that only for words the direct route to the mental lexicon is successful in mapping the orthographic percept onto stored word form representations.

**Left MTG**

Regions of the left lateral temporal lobe are long known to be involved in language processing. The posterior aspects of the left superior and middle temporal lobe correspond to Wernicke’s area. In the present study, the posterior part of the left MTG was activated reliably stronger for words than for pseudowords. It has been suggested (e.g., Beauregard et al., 1997; Howard et al., 1992) that posterior MTG might be an area where visual word forms are stored. Citing reports of posterior MTG activation in studies that did not involve word reading, Cohen et al. (2000), who observed activation for words in the same region, argued convincingly that this area is unlikely to be associated with the visual word form system. Hagoort et al. (1999) also reported increased activation for visually presented words as opposed to pseudowords in the MTG, approximately 2 cm anterior to the activation maximum in the present study. They associated the left MTG with the activation of phonological word forms (cf. also Price et al., 1994). A similar conclusion was reached by Indefrey and Levelt (2000) in a meta-analysis of word production tasks. These authors concluded that in posterior segments of middle and superior temporal gyri, phonological word form representations are stored, which are drawn upon both in production and comprehension.

In other studies, activations of the MTG were associated with lexical–semantic processing (e.g., Pugh et al., 1996). An involvement of posterior temporal areas (i.e., Wernicke’s area) in lexical–semantic processing is also suggested by clinical psycholinguistic studies. It has been demonstrated that Wernicke’s aphasics are unable to semantically categorize words (e.g., Zurif, Caramazza, Myerson, & Galvin, 1974) or to explicitly judge words on the basis of semantic information (e.g., Goodglass & Baker, 1976), leading to the conclusion that controlled lexical–semantic processes are deficient in these patients (Milberg, Blumstein, & Dworetzky, 1987; see Friederici, 1998, for a review). However, the present data do not allow to dissociate these two possible roles of the posterior left MTG (i.e., phonological vs. semantic representations; cf. also Cohen et al., 2000).

**Left IFG and Anterior Insula**

As in many other studies of word processing, activation differences were also observed in the IFG of the left hemisphere. Left IFG activations are often associated...
with phonological processing, phonological retrieval, or phonemic analysis (for reviews, see Indefrey & Levelt, 2000; Fiez & Petersen, 1998; Friederici, 1998). In the present experiment, the IFG was not activated by manipulating semantic or phonological task demands directly. Activation was enhanced when the presumed search times in the mental lexicon (as indicated by decision times) were prolonged (i.e., in low-frequency words and in pseudowords relative to high-frequency words). Increased activation of superior left pars opercularis (BA 44) for low-frequency words and for pseudowords suggests that lexical search is mediated more strongly by phonological information in these items. Grapheme-to-phoneme conversion rules assemble the phonological form of the orthographic stimulus in the case that the correct lexical entry is not rapidly identified on the basis of visual word form information. As this is exactly what is predicted by the dual-route model, the present data can be taken to suggest that the assembled route of word processing, subserving grapheme-to-phoneme conversion, involves the superior division of the pars opercularis of the left IFG. This interpretation converges with the conclusions drawn by Fiez et al. from word naming (i.e., Fiez & Petersen, 1998; Fiez et al., 1999). However, our data reach beyond these previous reports by providing evidence that phonological information derived via the grapheme-to-phoneme conversion route is mediating lexical access, even when phonological codes do not have to be activated for the actual pronunciation of words.

The anterior insula, as well as portions of the thalamus and of the head of the caudate nucleus bilaterally, was also activated more strongly when low-frequency words and pseudowords were processed. It can therefore be concluded that these areas, together with left BA 44, constitute a network supporting grapheme-to-phoneme conversion during visual word processing. This finding demonstrates that the function of the insular cortex in language goes beyond speech articulation (e.g., Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Dronkers, 1996) and supports similar conclusions drawn by Fiez et al. (1999) regarding anterior insula activity during word naming. Unlike proposed by Indefrey and Levelt (2000), there was no evidence in the present study that activation of the anterior insula was associated with visual processing as all stimulus conditions were matched for number of letters and number of syllables.

Activation of the more anteriorly located pars triangularis of the left IFG (i.e., BA 45) in the contrast between words of low and high frequency was nonoverlapping with the activation in BA 44. Furthermore, this activation effect was not observable when pseudowords were contrasted with high-frequency words. Thus, while BA 44, the anterior insula, and subcortical structures were activated by lexical and nonlexical stimuli alike (i.e., by low-frequency words and by pseudowords), BA 45 activation was dependent on the lexical status of the stimuli. Therefore, we conclude that the two IFG activations are probably associated with different processes, although the present data do not allow to dissociate them directly. The more anterior portions of left IFG are generally associated with strategic processes of semantic retrieval (e.g., Friederici, Optiz, & von Cramon, 2000; Gabrieli, Poldrack, & Desmond, 1998; Fiez, 1997). It has been suggested that the IFG does not subserve semantic retrieval per se but is activated when the task requires selection of information from among competing alternatives in semantic knowledge (Thompson-Schill, D’Esposito, Aguirre, & Farah, 1997). We suggest that the functionality of the anterior portions of the left IFG can be extended to incorporate lexical–semantic selection processes. Most models of word processing in the visual domain, like the search model proposed by Forster (e.g., Bradley and Forster, 1987; Forster, 1976), assume that during the very first stages of lexical access, several competing candidate entries are activated on the basis of initial visual access codes, which do not necessarily represent the full visual input. Full lexical access requires that the best-matching entry has to be selected from the preactivated candidate set. It might be hypothesized that the strategic control of lexical selection from among competing preactivated entries is associated with stronger activation in left pars triangularis. This proposal, however, needs to be investigated further.

Conclusions

Through the use of event-related fMRI and the manipulation of stimulus properties in the lexical decision task, we were able to provide evidence for distinct functional contributions of different brain regions to lexical access. A bilateral occipito-temporal network is involved in the prelexical recognition of visual word forms. This network includes the fusiform gyrus, where it is assumed to provide the access code to higher-level semantic or phonological information represented in the posterior left MTG. Neural activity in the left IFG, the anterior insula, and in subcortical structures was influenced by word frequency, suggesting that these brain regions come into play when lexical access is difficult. The superior pars opercularis (BA 44), together with the anterior insula and subcortical structures, appears to mediate lexical access through grapheme-to-phoneme conversion, while the pars triangularis (BA 45) might support mechanisms of selection from the mental lexicon.

METHODS

fMRI Data Acquisition

Twelve axial slices (5 mm thickness, 2 mm interslice distance, FOV 19.2 cm, data matrix of 64 × 64 voxels,
acquisition bandwidth 100 kHz, voxel size $3 \times 3 \times 5$ mm) were acquired every 1.5 sec during functional measurements (BOLD sensitive gradient EPI sequence, TR = 1.5 sec, TE = 30 msec) with a 3-T Bruker Medspec 30/100 system. The upper border of the fifth slice was aligned to the AC–PC line. Prior to functional imaging, T1-weighted MDEFT images (data matrix $256 \times 256$, TR = 1.5 sec, TE = 10 msec) were obtained with a non-slice-selective inversion pulse followed by a single excitation of each slice (Norris, 2000) for coregistration of functional scans with previously obtained high-resolution whole-head 3-D MDEFT brain scans (128 sagittal slices, 1.5 mm thickness, data matrix $256 \times 256$ voxels; cf. Lee et al., 1995). T1-weighted structural EPI images (TI = 1200 msec, TR = 30 sec, TE = 30 msec) were also acquired to allow for a refinement of the coregistration of functional data sets to the high-resolution structural scans.

Participants

Twelve paid volunteers (five women; 20–31 years; mean age 24.8 years) participated in the study. All participants were students of the University of Leipzig and without a history of neurological or psychiatric disorders. All participants had normal or corrected-to-normal vision, were native speakers of German, and were right-handed (laterality quotients of 100% except for one participant who had a laterality quotient of 91% according to the Edinburgh handedness scale; Oldfield, 1971).

Material

Word stimuli were one- or two-syllabic inanimate nouns with four to seven letters. They consisted of 68 low-frequency words that had frequencies of either zero or one in one million (mean frequency 0.5) and 68 high-frequency words with a mean frequency of 149.75, as determined on the basis of the Mannheim written and spoken corpus of the CELEX database (http://www.kun.nl/celex/). High- and low-frequency words were matched by constructing word pairs that had the same initial letter, number of syllables, and number of letters. The 68 word pairs used in the experiment were selected from a pool of 86 pairs on the basis of a behavioral pretest ($n = 10$). In addition, the selected words were rated for concreteness by eight individuals. Words of high and low frequency had comparable levels of concreteness. There was a slight but nonsignificant tendency for low-frequency words to be more concrete than high-frequency words [$F(1,7) = 4.2; p = .08$]. A total of 136 pseudowords were constructed from the word stimuli by randomly exchanging one or two letters (but not the initial letter). Ten alternative pseudowords were generated for each of the word stimuli and one phonologically plausible candidate was selected.

Experimental Procedure

During the experiment, 272 (i.e., $68 \times 2 \times 2$) trials of 7.5-sec length (i.e., five scans) were presented to the participants. At the beginning of each trial, a cue consisting of three asterisks in the center of the screen was presented for 100 msec. After another 400 msec, the item was shown for 400 msec in light-gray letters on a dark-blue screen. Consistent with the German orthography, all stimuli were presented with an initial uppercase letter and the rest of the word printed in lowercase. Participants were instructed to indicate as fast as possible whether the presented stimulus was a legal German word or not by pressing buttons on a response box (i.e., to make a lexical decision). After the response, but at least 400 msec after stimulus offset, a feedback was given visually. Thus, the minimum onset latency of the feedback cue was 800 msec after the onset of the item. Maximal response time allowed was 2 sec. The effective ISI between the offset of an item and the onset of the next item was 7.1 sec.

An experimental session consisted of four runs of 8 min 45 sec in which low- and high-frequency words and pseudowords were presented in pseudorandomized order. Two differently randomized stimulus sequences were constructed for the experiment. In these lists, words and pseudowords, as well as low- and high-frequency words, were distributed evenly across the four runs to assure that the same number of items of each condition was presented in each run. No more than three items of the same type (i.e., words or pseudowords) could be presented in a row and the frequency of transitions between the different event types was balanced such that each event type was followed equally often by each of the other event types. Participants were randomly assigned to one of the two lists.

Data Analysis

Data processing was performed using the LIPSIA software package (Lohmann et al., in press). Preprocessing steps included a slice time correction using sinc interpolation, a motion correction (rigid-body realignment), and spatial smoothing using a Gaussian kernel with 5.6 mm FWHM. Statistical analyses were performed using procedures of the SPM package (Josephs, Turner, & Friston, 1997; Friston et al., 1995) and an event-related design in which only correctly answered trials were included. The mean number of errors per event type is given in Table 1. The balancing of the design was not critically affected by the exclusion of incorrect trials. The first 10 images in each data set were excluded from the analysis. Observed data and design matrix were convolved by a Gaussian kernel with 4 sec FWHM. High pass filtering was performed during parameter estimation using the general linear model. Cutoff frequencies were determined on the basis of frequency and distribution of
the four event types within the design matrix (Friston, 1999). For each participant, parameter estimates were stored in individual contrast images that were registered and normalized to stereotactic space (Talairach & Tournoux, 1988) and subsequently transformed nonlinearily to match a reference brain (Thirion, 1998). Normalized contrast images were postprocessed in a second-level analysis using a one-sample t test for all subjects (Holmes & Friston, 1998). Resulting statistical parametric maps (i.e., SPM{Holmes & Friston, 1998}) were thresholded at 3.09 (i.e., \( p = .001 \), one-tailed, uncorrected).

To control for multiple statistical comparisons, corrected significance values were computed at the cluster level using distributional approximations from the theory of Gaussian fields. The smoothness of the activation map was calculated during parameter estimation of the general linear model on the basis of the standardized residuals (Kiebel, Poline, Friston, Holmes, & Worsley, 2000). Estimated smoothness and size of the search volume (i.e., number of resels) of the different contrasts are given in the respective figure captions. Cluster-size probabilities (i.e., the probability of obtaining a cluster of the given size or larger) were calculated on the basis of the number of resels in the functional data (Worsley et al., 1996). A spatial extent threshold of 200 mm\(^3\) was used for the identification of reliably activated clusters. Only activation clusters that were significant at the cluster level (i.e., \( p < .05 \)) are reported.

**Acknowledgments**

The authors wish to thank Thom Gunter, Sonja Kotz, Axel Mecklinger, Edward Smith, David Swinney, and two anonymous reviewers for their comments on earlier versions of this paper. Thanks are also due to Anke Mempel, Mandy Naumann, and Kathrin Wiesner for help with data acquisition. This work was supported by the Leibniz Science Price awarded to A.D.F. by the German Research Foundation.

Reprint requests should be sent to Christian J. Fiebach, Max Planck Institute of Cognitive Neuroscience, PO Box 500 355, D-04303 Leipzig, Germany, or via e-mail: fiebach@cns.mpg.de.

The data reported in this experiment have been deposited in The fMRI Data Center (http://www.fmridc.org). The accession number is 2-2001-111XN.

**Note**

1. Generally, the opposite relation between word frequency and concreteness is observed. Low-frequency words are known to be more abstract than high-frequency words. Accordingly, this tendency does not confound the contrast between high- and low-frequency words but works in the opposite direction, if it has an effect at all. The clear behavioral and activation effects elicited by word frequency, however, make this possibility very unlikely.

**REFERENCES**


