

A Neural Signature of Phonological Access: Distinguishing the Effects of Word Frequency from Familiarity and Length in Overt Picture Naming

William W. Graves, Thomas J. Grabowski, Sonya Mehta,
and Jean K. Gordon

Abstract

■ Cognitive models of word production correlate the word frequency effect (i.e., the fact that words which appear with less frequency take longer to produce) with an increased processing cost to activate the whole-word (lexical) phonological representation. We performed functional magnetic resonance imaging (fMRI) while subjects produced overt naming responses to photographs of animals and manipulable objects that had high name agreement but were of varying frequency, with the purpose of identifying neural structures participating specifically in activating whole-word phonological representations, as opposed to activating lexical semantic representations or articulatory-motor routines. Blood oxygen level-dependent responses were analyzed using a parametric approach based on the frequency with which each word produced appears in the language. Parallel analyses were performed for concept familiarity and word length, which provided indices of semantic and articulatory loads. These analyses permitted us to identify regions related to word fre-

quency alone, and therefore, likely to be related specifically to activation of phonological word forms. We hypothesized that the increased processing cost of producing lower-frequency words would correlate with activation of the left posterior inferotemporal (IT) cortex, the left posterior superior temporal gyrus (pSTG), and the left inferior frontal gyrus (IFG). Scan-time response latencies demonstrated the expected word frequency effect. Analysis of the fMRI data revealed that activity in the pSTG was modulated by frequency but not word length or concept familiarity. In contrast, parts of IT and IFG demonstrated conjoint frequency and familiarity effects, and parts of both primary motor regions demonstrated conjoint effects of frequency and word length. The results are consistent with a model of word production in which lexical-semantic and lexical-phonological information are accessed by overlapping neural systems within posterior and anterior language-related cortices, with pSTG specifically involved in accessing lexical phonology. ■

INTRODUCTION

Language processes acting on word-level representations, referred to as *lexical* processes, were among the first cognitive phenomena to be studied with functional neuroimaging (e.g., Petersen, Fox, Posner, Mintun, & Raichle, 1988). The investigation of these processes has benefited greatly from the availability of detailed cognitive models and the empirical evidence which supports them. These models have allowed interpretation of certain behavioral effects in terms of the cognitive processing stages at which they arise. In turn, the neural correlates of these behavioral effects have been investigated with functional imaging to help delineate neural systems supporting specific processing stages (e.g., Mechelli et al., 2005; Binder et al., 2003). In this study, we explore the physiologic correlates of the word frequency effect (WFE) in picture naming in order to delineate neural systems involved in accessing lexical phonology.

A challenge facing any study in which a specific psycholinguistic variable is used to probe a particular level of language processing is dealing with the degree to which the independent variable of interest may be correlated with other variables not associated with the level of processing being investigated. Word frequency, for example, is potentially correlated with the familiarity of the concepts the words denote, or the length of the words themselves. To distinguish areas specifically involved in lexical phonological access from those which interact with neighboring processes such as lexical semantic or articulatory-motor program access, we performed parallel parametric analyses to delineate areas significantly associated with producing either words of decreasing concept familiarity or increasing length (in terms of number of syllables) from those exhibiting a significant effect of decreasing word frequency.

Models of Lexical Access

Models of lexical access during speech production generally contain two lexical processing stages, one in which

University of Iowa

the meaning of a word (lexical semantics) is accessed, and another in which its sound code (lexical phonology) is accessed (Rapp & Goldrick, 2006; Levelt, Roelofs, & Meyer, 1999; Griffin & Bock, 1998; Caramazza, 1997; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997). Within lexical semantics, some investigators have also proposed an amodal intermediary representation known as the lemma, containing lexical semantic and some syntactic information but occurring prior to lexical phonological access (Levelt et al., 1999; Levelt, 1989). An example of such a two-stage model for picture naming is shown in Figure 1. The picture naming process leads in with visual perception, recognition, and accession of nonverbal conceptual knowledge of the depicted entity. Following this, there are two stages of lexical access: First, a lexical semantic representation is accessed, then a lexical phonological representation (the phonological word form, or lexeme).

Models differ as to the degree of discreteness or seriality of information flow between stages. Whereas strictly serial models (e.g., Levelt et al., 1999) hold that

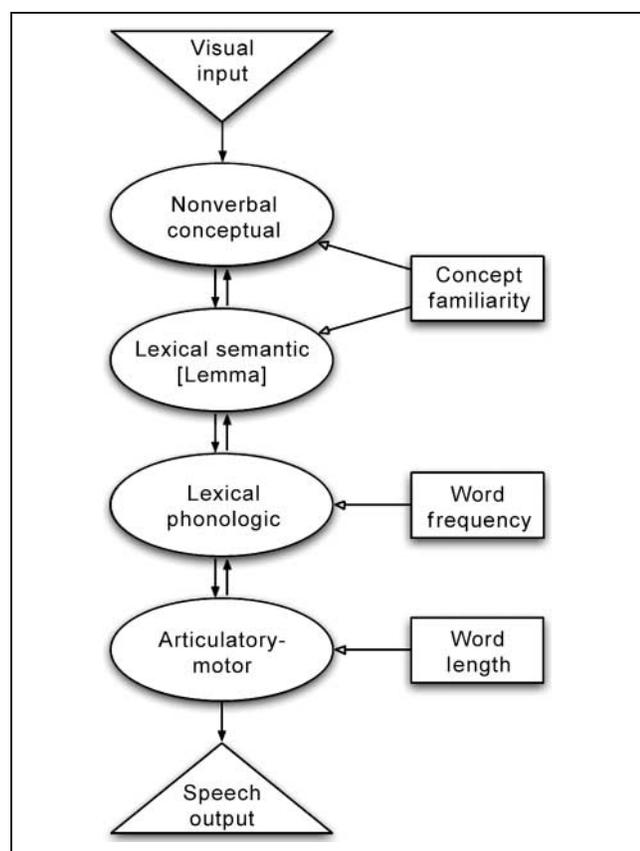


Figure 1. Processing model of word production in picture naming. Ovals represent major representational stages in the picture naming process. Rectangles with open arrows represent factors thought to modulate the cognitive resources used to access the indicated representations, and which are analyzed parametrically in the current experiment. Triangles represent input/output modalities. Further details are provided in the text.

lexical semantic access must be completed before lexical phonological access can begin to take place, cascade models (e.g., Humphreys, Riddoch, & Quinlan, 1988) propose that lexical phonologic access can begin before lexical semantic access is complete, and interactive accounts (e.g., Dell et al., 1997) propose that processing at the lexical phonological level can feed back to the lexical semantic level. Although these models differ in their specific accounts of information flow dynamics, there is a general consensus that, at least for picture naming, lexical semantic information normally begins to be accessed prior to accessing lexical phonology (Rapp & Goldrick, 2006; Levelt et al., 1999; Caramazza, 1997; Dell et al., 1997; Levelt et al., 1991; Schriefers, Meyer, & Levelt, 1990). A variable thought to modulate the resources required to access the lexical phonological information corresponding to the lexical semantic representation is word frequency.

The Word Frequency Effect

The WFE in picture naming was first characterized by Oldfield and Wingfield (1965), and has since proven to be a highly replicable and reliable effect (e.g., Székely et al., 2003; Snodgrass & Yuditsky, 1996; Jescheniak & Levelt, 1994). They found that the time interval between the presentation of the stimulus and the onset of the naming response was significantly predicted by the frequency of the response word, with low-frequency words producing comparatively longer response delays. Their observation left open the question of whether the WFE was the result of, for example, lack of familiarity with what the less frequent words represented (i.e., of semantic origin), or alternatively, the result of a lack of practice with producing the less frequent words (i.e., of lexical phonological origin). Evidence that the WFE acts at the level of lexical phonological access, rather than lexical semantic access, has come from the frequency inheritance effect, in which word production latencies for low-frequency homophones are affected by their high-frequency homophone counterpart. For example, NUN, a lower-frequency word, is produced with speed comparable to its higher-frequency homophone, NONE. The same is true, for example, for MOOR and MORE (Jescheniak, Meyer, & Levelt, 2003; Jescheniak & Levelt, 1994). Additionally, subjects show reduced error rates for low-frequency homophones (BEE) that match those of their high-frequency counterparts (BE) (Dell, 1990), even though low-frequency words are generally more vulnerable to phonological speech errors (Stemberger & MacWhinney, 1986). Thus, homophones, despite having separate meanings, act as if they share a single phonological representation that combines their individual frequencies, benefiting the low-frequency homophone in production tasks. The frequency inheritance effect is not seen in picture verification tasks, indicating that the effect arises at the lexical, rather than at the conceptual,

level (Jescheniak & Levelt, 1994; Morrison, Ellis, & Quinlan, 1992; Wingfield, 1968). Finally, execution of articulatory movements to pronounce a word does not consistently give rise to frequency effects (Jescheniak & Levelt, 1994; Monsell, 1990; Forster & Chambers, 1973). Taken together, these findings suggest that the WFE acts primarily at the level of lexical phonological access during word production.

It should be noted, however, that locating the WFE specifically at the level of phonological access is not without controversy, even among models that propose two stages of lexical access. Several computational models that feature interactive processing between lexical levels show that when a WFE emerges from training or parameter tuning to fit a sample dataset, the effect is underlain by changes in weights throughout the network (Harm & Seidenberg, 2004; Dell et al., 1997; Plaut, 1997). Interactive models challenge the assumption that word frequency acts exclusively at the level of whole-word phonological access. Interactivity of information flow among representational levels seems neurally realistic, and models incorporating interactivity have successfully accounted for patterns of word production errors seen in both unimpaired and brain-damaged subjects (for a recent review, see Rapp & Goldrick, 2006). It is for these reasons that we use double arrows in Figure 1 to indicate the dynamic nature of processing assumed in this word production model. However, interactivity is compatible with functional specialization, and it is likely that the cognitive resource costs associated with the WFE correspond most conspicuously to neural work (indexed by blood oxygen level-dependent [BOLD] signal in magnetic resonance imaging [MRI]) in the portions of the network specialized for accessing word-level phonological representations.

A second practical issue is that word frequency is correlated with concept familiarity and word length, with lower-frequency words tending to be less familiar and longer. These correlations are difficult to avoid in large stimulus sets representing concrete entities, and such correlations are potentially problematic for the specificity of interpreting WFEs. For example, to the extent that word frequency is correlated with concept familiarity, effects attributed to the lexical phonological level may be confounded with lexical semantic effects. A similar problem arises from the correlation of word frequency with word length, leading to a possible confounding of lexical phonological level effects with those more related to articulatory-motor programs. We overcome these confounds in the current imaging experiment by performing parallel analyses of these parameters and comparing the anatomic loci of WFEs to those of familiarity and length effects. This approach enables detection of areas where frequency interacts with familiarity or length effects (and therefore may be confounded with them), and areas where activation is associated with the WFE alone.

Several functional imaging studies have utilized WFEs manifested in tasks other than picture naming, including lexical decision (Fiebach, Friederici, Muller, & Yves von Cramon, 2002), semantic decision (Chee, Hon, Caplan, Lee, & Goh, 2002), and silent word reading (Kronbichler et al., 2004). These studies demonstrate that WFEs are detectable with functional MRI (fMRI). However, those not concerned with word production likely address frequency effects that arise from different cognitive and neural processes than the present study. In fact, few imaging studies have specifically investigated the WFE in overt picture naming.

Neural Systems Predictions

To connect constructs from the cognitive model pictured in Figure 1 and described above with expected results from functional brain imaging data, bridging assumptions must be made about the relationship between these two kinds of information. The psycholinguistic variables considered in this study follow an approximately continuous distribution where their changing values correlate with changes in response latencies (e.g., decreasing word frequency correlated with increasing naming latencies). Here we manipulate psycholinguistic variables which have been shown to lead to processing time changes in chronometric studies. We then assume, as illustrated in our cognitive model, that this variance in processing time is the result of changes in processing demands for accessing the relevant level of representations. Such changes in processing demands will presumably correspond to changes in duration and intensity of physiologic work at synapses in neural regions that underlie the relevant cognitive operations. If these psycholinguistic variables are indeed affecting neurophysiologic processing in a manner commensurate with how they are thought to affect cognitive processing, their effects will appear in the functional imaging data as changes in BOLD signal. For example, comparatively low-frequency or familiarity values, which are associated with increased processing demands in the cognitive framework, and expressed as more processing time in chronometric studies, will appear as increases in BOLD signal in regions involved in accessing/activating lexical phonological and lexical semantic representations, respectively.

With this foundation in place, it remains to specify a framework for interpreting data at the neural systems level that is comparable to the specified cognitive framework. Damasio and colleagues have proposed a neural systems level framework which accounts for cognitive phenomena such as word and concept retrieval (Tranel, Grabowski, Lyon, & Damasio, 2005; Damasio, Tranel, Grabowski, Adolphs, & Damasio, 2004; Tranel, Damasio, & Damasio, 1997, 1998; Damasio, Grabowski, Tranel, Hichwa, & Damasio, 1996; Damasio & Damasio, 1994; Damasio, 1989a, 1989b). According to this framework,

recall is accomplished when sensory information stored in early sensory cortices is reactivated and bound with information from other sensory cortices. This process is directed by convergence zones located in heteromodal association cortices. In the context of lexical retrieval, the relevant putative convergence zones are those that bind information in two distinct neural systems: those supporting semantic knowledge, which involve posterior cortical and possibly prefrontal regions, and those supporting language implementation, in the left perisylvian region (Damasio et al., 1996, 2004; cf. Levelt, 1999). Such convergence zones would participate in the access of lexical semantic and lexical phonological representations, and, if concentrated in specific anatomical locations (i.e., convergence *regions*), would be detectable with fMRI.

Regarding the specific neural systems hypothesized to be involved in lexical phonological access for word production, two areas in the left hemisphere bordering the sylvian fissure, the inferior frontal gyrus (IFG) and posterior superior temporal gyrus (pSTG), have been consistently shown to support word production (as reviewed, for example, by Alexander, 2003; Saffran, 2000). The left IFG is known to be critical for speech production, but its exact role in lexical phonological access is less clear. There are at least two major competing hypotheses. The first holds that lexical processing in the left IFG is domain-specific, with lexical phonology processed in its more dorsal/posterior regions (BA 44/45), and lexical semantics processed in more ventral/anterior ones (BA 47/45) (Bookheimer, 2002; Poldrack et al., 1999; Fiez, 1997). According to an alternate account, the left IFG supports domain-general processes characterized by either response selection (Thompson-Schill, Bedny, & Goldberg, 2005; Thompson-Schill, 2003), controlled recall (Badre & Wagner, 2002; Gold & Buckner, 2002), or possibly both (Moss et al., 2005). Response selection involves picking among multiple competing alternatives in order to produce an appropriate response, and might occur in picture naming if, for example, a picture of a dog evokes not only DOG, the basic-level target name, but also RETRIEVER, or ROVER. The role of controlled retrieval in picture naming is less clear. Indeed, for at least one study (Moss et al., 2005), picture naming was chosen because it was thought to minimize demands on controlled retrieval.

Although these studies provide an entry point for conceptualizing IFG function in lexical retrieval, it should be noted that most, if not all, of these studies employed tasks which did not require overt speech responses. Further, their general use of cognitive subtraction techniques, in contrast to the parametric approach employed here, may also limit the extent to which findings from these previous studies may be used to predict results from the current one. To the extent that comparisons can be drawn, there are at least four possible outcomes that would speak to the role of the IFG in lexical retrieval: (1)

WFE-specific activation in the dorsal/posterior IFG, (2) WFE-specific activation in the anterior/ventral IFG, (3) the presence of both word frequency and concept familiarity effects in the dorsal/posterior IFG, and (4) the same conjunction of effects in the anterior/ventral IFG. Assuming that the WFE is predominantly related to lexical phonological access, the domain-specific account would predict Outcome 1, whereas a domain-general view would predict Outcomes 3 and 4. We will return to this point in the discussion.

An alternate account of IFG function, which is not necessarily incompatible with that described above, comes from the prefunctional imaging era. Specifically, the pars opercularis and triangularis of the IFG constitute Broca's area, which has traditionally been associated with the motor implementation of speech (Broca, 2006). The pars opercularis (roughly BA 44) adjoins the primary motor cortex, and the inferior motor cortex is an area identified by Indefrey and Levelt (2004) as being associated with articulatory programming in their meta-analysis of word-production studies. Thus, the primary motor cortex and/or the pars opercularis are expected to exhibit effects of word length.

The left pSTG (corresponding to Wernicke's area) has been implicated in speech production generally (Saffran, 2000). Anatomically, it is connected to the IFG and the inferotemporal cortex (IT) through the arcuate fasciculus, possibly through an anterior segment connecting the pSTG and the IFG and a posterior segment connecting the pSTG and the IT (Catani, Jones, & ffytche, 2005). The pSTG is considered multimodal association cortex in that it receives information from both the auditory and visual association cortices (Mesulam, 1998). As predicted by classical lesion evidence, functional imaging has shown the pSTG to be more specifically related to production and comprehension of speech than the more anterior sectors of the STG (Wise et al., 2001). Evidence for its specific involvement in lexical phonological access was assembled by Indefrey and Levelt (2000, 2004) in two broad-scoped meta-analyses of word production imaging studies. Both meta-analyses pointed to a specific role for the left pSTG in lexical phonological access.

Evidence has also accrued pointing to a role for the left IT in some aspects of language processing. For example, the left IT (of which the most posterior aspect bordering the occipital lobe is referred to as the occipitotemporal cortex, or OT) is one of the most consistently activated areas during picture naming in normal subjects (Murtha, Chertkow, Beaugard, & Evans, 1999). In addition, damage to the left IT has been associated with impaired naming ability in the context of otherwise fluent speech (Damasio et al., 1996, 2004; Raymer et al., 1997; Coughlan & Warrington, 1978). The observation that the patients with damage to the left IT in the Damasio studies were able to describe the entity to be named when they were unable to produce the name

itself, together with a generally fluent pattern of speech, suggests that there may be an area in the left IT that acts to link conceptual information to word production rather than support conceptual or word representations per se.

In summary, the current study aims to use the WFE to selectively engage the neural areas participating in lexical phonological access, and to further distinguish this level of processing from that of lexical semantic or articulatory-motor access by comparing the WFE with that of familiarity and length. We expected to demonstrate a WFE for word production during picture naming, and hypothesized that parallel parametric analyses would reveal WFE-specific neural activity in the left hemisphere IFG, pSTG, and IT.

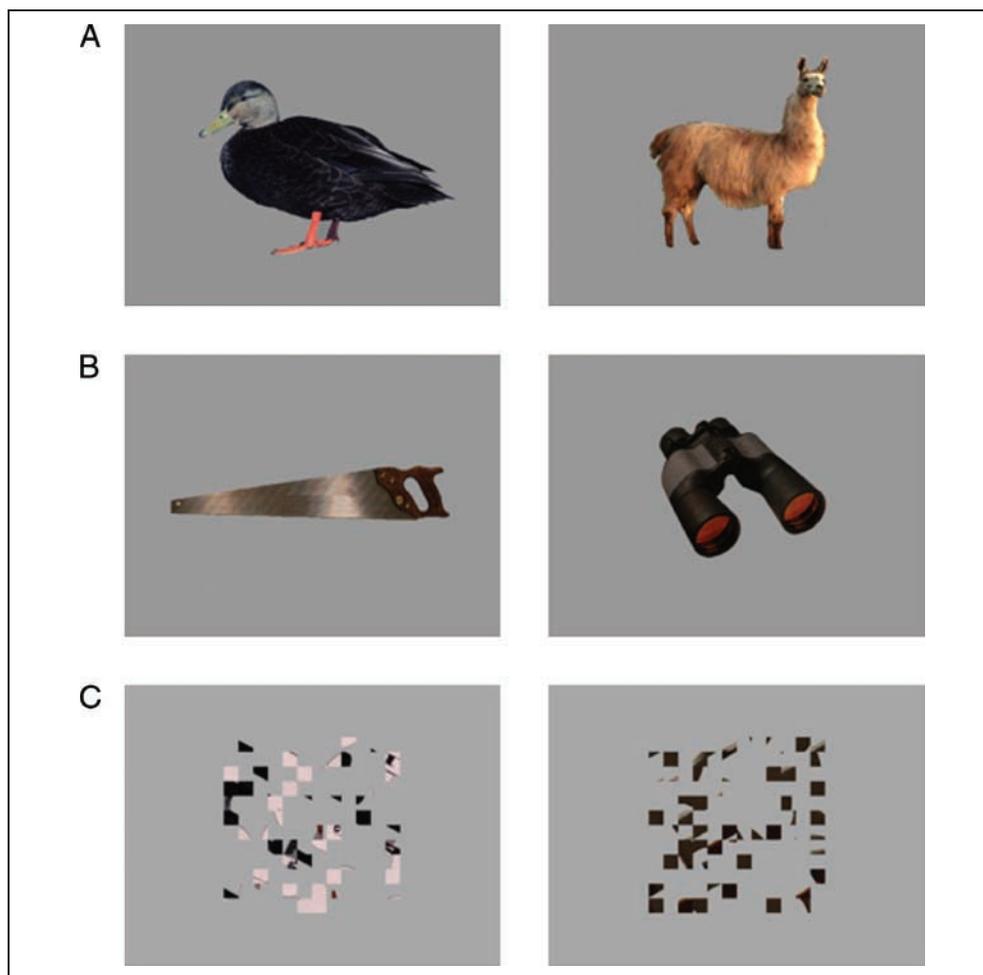
METHODS

Stimulus Norming

Fifty-nine neurologically normal undergraduate students from the University of Iowa took part in stimulus rating sessions for course credit. Normative data were obtained on 95 color photographs of animals and 101 color photo-

graphs of manipulable objects (tools). These two categories were chosen to enable the study of category effects during naming, which are beyond the scope of the current report. For the purposes of gathering normative ratings, stimuli were presented in a classroom environment using Microsoft PowerPoint 2002. Subjects were asked to name each image and rate them for familiarity and image agreement (following Fiez & Tranel, 1997). Naming responses for each stimulus were analyzed for level of agreement by calculating an information statistic, H , based on Snodgrass and Vanderwart (1980). Familiarity and image agreement ratings were based on a 5-point scale, with 5 being high familiarity and image agreement, respectively. Items with low name agreement values ($H > 1.5$), low image agreement ratings (< 4), or low familiarity ratings (< 3) were excluded from the fMRI sessions, leaving a total of 156 unique stimuli (82 animals and 74 manipulable objects, some examples of which are given in Figure 2). The set of stimuli used in fMRI had a mean name agreement of $H = 0.51 \pm 0.43$ (low values correspond to better name agreement), a mean image agreement of 4.25 ± 0.48 , and a mean familiarity of 4.26 ± 0.45 . These values generally correspond to those seen in the picture naming literature.

Figure 2. Examples of photographic stimuli for naming nonunique entities. (A) Animals: duck (relatively high frequency), llama (relatively low frequency); (B) Manipulable objects (tools): saw (relatively high frequency), binoculars (relatively low frequency); (C) Tiled stimuli (left: light, right: dark) used in the light/dark comparison task included to allow for imaging the main effect of naming but not analyzed for the purposes of this study.



The visual complexity of each stimulus was considered to be equivalent to its compressed electronic file size (mean 134.51 ± 55.95 kb) (Székely & Bates, 2000). Word frequencies (number of occurrences per million words) and phonological neighborhood densities (number of words sharing all but one phoneme with the target) for each correct response given at scan time were obtained from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). The mean phonological neighborhood density (defined as the number of words differing from the target word by the substitution of a single phoneme) was 10.18 ± 29.10 . This definition of phonological neighborhood density is restrictive compared to definitions which include not only substitutions but also additions and deletions as neighbors (e.g., Gordon, 2002; Vitevitch, 2002). Here we chose substitutions only in order to respect CVC structure (Shattuck-Hufnagel, 1992).

For word frequencies, lexeme frequencies (in number of occurrences per million) were used instead of lemma frequencies because we were interested in the frequencies of word forms rather than, for example, separate frequencies for appearances of nouns and verbs. The negative base-10 logarithm of the lexeme frequency of each response word was then calculated according to the formula $f(x) = -\log_{10}(x)$, where x is the number of occurrences per million from CELEX. This expression was used so that frequency would correlate linearly with response latency, and presumably, with the BOLD effect as well. By this calculation, lower-frequency words end up having higher numerical values. For the responses obtained during scanning, the mean $-\log$ frequency was 4.21 ± 0.86 . This is on the order of that obtained by Jescheniak and Levelt (1994) for Dutch words in CELEX and by Székely et al. (2003) for English words in CELEX (4.11 and 4.82, respectively, after applying the negative log transform). There was no significant difference in word frequency for actual scan-time naming responses to animal photographs compared to tool photographs (animal: 4.27 ± 0.72 ; tool: 4.15 ± 0.98 ; $p = .38$).

Subjects

Subjects were recruited for fMRI scanning from the University of Iowa community and were paid \$50 for participation in the 2-hour scanning session, which is in compliance with local institutional review board guidelines. Inclusionary criteria required that all subjects be right-handed (a score of at least +85 on the Oldfield–Geschwind handedness questionnaire), have no history of neurologic or psychiatric disease, and be able to undergo an MRI scan. There were 14 subjects (11 women), with a mean age of 24.1 ± 5.7 years.

Task

Subjects were informed that they would be shown pictures of “animals and everyday objects,” and to “name

each picture as quickly and accurately as possible” (naming condition). If they felt that they did not know the name of the pictured entity, they were to say “pass.” Subjects were also told that randomly interspersed with these pictures would be color photographs scrambled into multiple tiled pieces. For these, subjects were to respond “light” if the pieces were mostly light hued, or “dark” if the pieces were mostly dark hued (light/dark condition). Examples of stimuli for the naming condition and the light/dark comparison condition are given in Figure 2, and additional examples are provided by Tranel et al. (2005). The light/dark task is treated as a covariable of no interest and plays no further role in the analyses reported here. All responses were spoken overtly.

Scanning Session

Images were acquired with a General Electric LX CV/i scanner at a field strength of 1.5 T using a transmit-and-receive quadrature head coil. Time-series images were acquired using a T2* gradient-echo single-shot echo-planar image (EPI) sequence (TE = 40 msec, TR = 2 msec, FOV 24 cm, matrix 64×64). For each run, 220 time points, composed of either 20 or 24 contiguous oblique axial slices (5 mm thick), were acquired parallel to the intercommissural plane and covering the whole brain. An 8-shot echo-planar T2*-weighted volume (matrix 128×128) and T1-weighted acquisition (SPGR, flip angle = 30° , TR = 24 msec, TE = 7 msec, FOV = 24 cm, matrix 256×192) were also acquired in the same oblique axial orientation as the EPI time series. Finally, a high-resolution anatomical scan (SPGR, flip angle = 30° , TR = 24 msec, TE = 7 msec, $256 \times 192 \times 124$, slice thickness = 1.5–1.6 mm) was obtained.

Overt verbal responses were obtained using a microphone attached to the distal end of a pneumatic tube secured onto the head coil near the subject’s mouth. Stimuli were delivered through a video projector aimed at a rear projection screen secured to the end of the scanner bed near the subject’s feet. Subjects viewed the projected stimuli through mirrors attached to the head coil. The stimulus display essentially lay within the foveal field, subtending approximately 3.7° of visual angle. All data acquisition and stimulus delivery events were time stamped with submillisecond resolution using the Input/Output time-aWare Architecture (I/OWA) system (Smyser, Grabowski, Frank, Haller, & Bolinger, 2001). Use of the I/OWA time-aware system obviates the need for explicitly synchronizing stimulus delivery and scanner TR, eliminates the need for slice-timing correction, and facilitates extraction of response latencies from stimulus and speech onset times (Grabowski et al., 2006; Mehta, Grabowski, Razavi, Eaton, & Bolinger, 2006).

Naming and light/dark stimuli were randomized together with a variable interstimulus interval (ISI; mean

3.93 sec, range: 1.51–20.48 sec). Including a wide range of ISIs allowed the mean ISI to be kept short while maintaining enough variance in the hemodynamic signal to detect stimulus-related changes in a rapid event-related design. Stimuli were displayed for 1000 msec. No stimuli were presented during the first or last 10 sec of each run.

Image Processing

All image registration operations were performed using Automated Image Registration, AIR 5.2.3 (Woods, Grafton, Holmes, Cherry, & Mazziotta, 1998; Woods, Grafton, Watson, Sicotte, & Mazziotta, 1998). The first three images in the time series were discarded to avoid saturation effects. Images within a run were aligned to the fourth image of the time series using a 3-D six-parameter rigid-body model. Data from all runs were then aligned to the average image of the first run and later analyzed in this orientation after smoothing with a Gaussian kernel ($7.5 \times 7.5 \times 10$ mm full width, half maximum). The geometric distortion between the single-shot EPI and the T1-weighted images of the same slices was corrected using spatial information from an eight-shot EPI (a more complete description of this process is provided by Mehta et al., 2006). Each subject's structural scan was registered to a Talairach-compatible atlas (Woods, Dapretto, Sicotte, Toga, & Mazziotta, 1999; Talairach & Tournoux, 1988) using a low-order non-linear warp. The derived transformations were later applied to the statistical images to allow for group analyses.

Speech Processing

In order to obtain latency data and judge the content of the responses, we used custom software implementing a time-aware spectral subtraction algorithm which removed noise due to scanner activity while leaving the speech signal intact (Mehta et al., 2006). Overt responses were automatically paired with stimulus presentation times and then checked manually for accuracy of assignments before final calculation of speech onset latencies. Naming latencies for successful responses were then correlated with negative log-transformed frequencies of the actual words produced during the scan. Successful responses were those which either matched the dominant responses given by the independent group of raters or had the same meaning. For example, “rabbit” or “bunny” were both considered successful responses (with the corresponding frequency of the actual utterance used in the analysis), but only “duck” and “llama” were accepted for the pictures in Figure 2A. Preverbal utterances, false starts, omissions, saying “pass,” providing an incorrect name (misidentifications), rushed responses where one name runs into the next, or response latencies more than three standard deviations above the mean (which could result either from abnor-

mally slow responses or a technical failure), were all considered unsuccessful responses (errors).

Image and Speech Analysis

fMRI time-series data were analyzed voxelwise using `tal_regress`, a custom software module that implements the general linear model (Frank, Damasio, & Grabowski, 1997). The regression model included three task-independent covariables for (1) global (mean) signal intensity, (2) a constant, and (3) four truncated Fourier series pairs with a cutoff of 1/110 Hz to model noise due to low-frequency drifts. The remainder of the covariables included in the regression model were each convolved with a hemodynamic response function and consisted of: (1) boxcar regressors for modeling the speech envelope; (2) the light/dark baseline task; (3) category (animal or tool); (4) a trial covariable to account for repeated exposures over time; (5) successful and (6) unsuccessful naming trials; and (7) either frequency, familiarity or length values nested within successful naming trials (as described further below). We did not analyze all three variables in one model because preliminary analyses of two subjects suggested the degree of multicollinearity among them resulted in inflated regression coefficients, making the approach problematic. Artifacts of overt speech were avoided by exploiting the hemodynamic lag to avoid colinearity of these artifacts with task regressors, and by the use of speech envelope regressors (Mehta et al., 2006).

All task covariables were generated from events defined to occur at stimulus onset times. These functions were used to analyze the data with voxelwise multiple linear regression, and the resulting regression coefficient images were transferred to a Talairach-compatible atlas space for group-level analysis. *t*-Statistic images were generated from the group regression images and thresholded at $p < .001$, before correcting for multiple comparisons. These steps were performed for each of the individual parametric analyses, where only the factor of interest differed. A spatial extent threshold of at least 11 contiguous voxels was subsequently applied to the uncorrected results of the parametric analyses in order to arrive at a corrected threshold of $p < .05$. This combination of uncorrected p value and spatial extent threshold is based on that derived by McDermott, Petersen, Watson, and Ojemann (2003), who used Monte Carlo simulations on random noise following a method described by Forman et al. (1995). The thresholded results of the parametric analyses were then used as input for the conjunction analyses, as described below.

The factors of interest for the parametric analyses were modeled as follows. For the word frequency data, a task reference function for successful naming was supplemented with an orthogonal covariate of interest defined by the deviation of the negative log-transformed word frequency of each response from the mean of the transformed word frequencies across the entire run. The

covariate of interest was orthogonal to the covariate for successful naming in that it indexed the WFE over and above the general effect of the naming task. This approach allowed production of words of different frequencies to be visualized with respect to each other, without the need for subtracting the effects of a baseline task (as mentioned earlier, the light/dark condition was simply entered as a covariable of no interest).

Two additional analyses of this type were performed, one for concept familiarity and the other for word length (number of syllables), each entered as a covariate in their respective models. To image the familiarity effect, ratings originally provided on a 5-point Likert-type scale (see above) were transformed as follows: $(-1 * x) + 5$, where x is the original familiarity rating. This flipped the scale so that less familiar concepts would be reflected by higher numbers. A familiarity covariable to explain BOLD signal variance was constructed by weighting the speech event times by the familiarity ratings and then convolving with a literature hemodynamic response function. In this way, overt responses to less familiar concepts were predicted to elicit increased hemodynamic response magnitudes compared to those rated as more familiar. A parallel analysis was carried out in a similar manner for word length effects. That is, producing words of more syllables was predicted to elicit larger-magnitude hemodynamic responses compared to producing words of fewer syllables.

Results of these three separate parametric analyses (one each for word frequency, familiarity, and length effects) were thresholded at $p < .05$, and binary masks of activated areas satisfying this threshold were then created. Clusters of voxels displaying significant effects of familiarity or word length were subtracted from clusters displaying significant frequency effects to yield a set of WFE-correlated activations with no overlap of familiarity or length effects (Figure 3.1). For the conjunction analysis, the masks for the word frequency and familiarity effects were multiplied so that any area, which was not significant in both, was reduced to zero, leaving only areas showing a significant conjunction. The same procedure was performed for word frequency and word length effects. Thus, the conjunction analysis used here corresponds to a logical AND, indicating effects of both conditions at the resulting locations (Nichols, Brett, Andersson, Wager, & Poline, 2005). For purposes of display, masks for each effect were assigned a different number (corresponding to a separate color) and overlaid together onto an averaged brain in Talairach-compatible atlas space (Figure 3.2).

RESULTS

Behavioral Measures

Mean latency to successful response onset was 1031 (± 160) msec. The frequency of the words the subjects produced when naming the pictures during scanning was significantly correlated with reaction time ($p < .01$), with

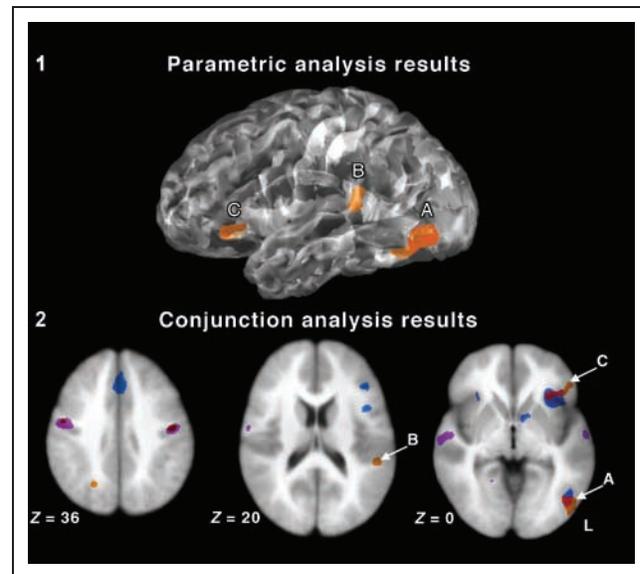


Figure 3. Results of (1) the parametric analysis of word frequency (after accounting for the effects of concept familiarity and word length) and (2) conjunction analyses. Row 1 is a three-dimensional rendering of the left hemisphere displaying word frequency-specific activations in orange, which are also shown in the axial slices in row 2 alongside concept familiarity (blue) and word length effects (purple). Letter labels for arrows on the axial slices correspond to those on the 3-D image, and denote the following regions: A, occipito-temporal cortex ($-47, -71, -3$); B, pSTG ($-51, -37, 20$); and C, IFG ($-43, 29, 0$). Note that areas A, B, and C show a WFE exclusive of, although in two cases adjacent to, familiarity and length effects. Also note that the 3-D representation shows only the three largest activations for the WFE derived from the parametric analyses. Orange = WFE; blue = familiarity effect; purple = length effect; red = frequency/familiarity overlap (right axial slice); dark red = frequency/length overlap (left axial slice at $z = 36$). For a full list of centers of activation for the parametric analyses isolating word frequency effects, see Table 2, and see Table 3 for conjunction analysis results.

reaction time changing 44.0 msec for each log unit of word frequency. The overall naming rate across subjects was 95.1%. For the remaining 4.9 % of trials, subjects either said “pass” (for 4.7% of trials, as instructed to do when they felt they could not produce the name) or subjects produced no response (0.2% of trials). Of the 95.1% of trials where subjects produced a name, content-related errors (12.6% of all trials, approximately half of which were misidentifications, e.g., seeing a picture of a deer and saying “moose,” or a picture of a level and saying “scale”) or producing a prenominal utterance (2.1% of all trials; e.g., saying “uh” before producing the name) occurred to give an overall “successful” response rate of 78%. “Rushed” responses, where one response ran into the next due to delayed response under time pressure, occurred for only 0.1% of all trials. Hence, “success” is defined strictly to mean not only success in content but also in execution. This success rate was comparable to that of Alario et al. (2004) for successful naming in their Experiments 1 and 2 (78.5%). An analysis by item in our study

revealed no significant difference in error rates across categories [$t(159) = 0.80, p = .43$]. Frequency was not significantly correlated with error rate ($r = .0026, p = .97$), which suggests that eliminating the unsuccessful items did not bias the estimation of frequency effects. Familiarity was significantly correlated with error rate ($r = -.29, p < .001$), indicating that pictures representing more familiar concepts were responded to more accurately overall than those representing less familiar concepts.

A regression analysis was performed to look at the effects of frequency, familiarity, and length on response latency. Visual complexity and phonological neighborhood density did not significantly predict response latencies, nor did they significantly correlate with frequency, and so will not be discussed further. The regression analysis revealed that both frequency and familiarity explained a significant, unique amount of variance in response latency ($\beta = 0.16, p < .05$ for frequency, and $\beta = -0.43, p < .0001$, respectively). Length did not significantly predict response latency but was significantly correlated with frequency ($r = .43, p < .0001$). Frequency and familiarity were also significantly correlated ($r = -.29, p < .001$).

The mean response latency reported here (1031 msec) is longer than the estimate (600 msec) provided by Indefrey and Levelt (2004). Although the longer response latencies obtained in the current study may have been affected by several factors (e.g., lack of pre-exposure to stimuli, use of color photographs instead of line drawings, or demands of the fMRI environment), the most likely explanation is the inclusion of a nonnaming baseline task in the event-related design. To the extent that the naming and light/dark conditions required different response preparation, it is likely that the longer response latencies were due to the random intermixing of naming and light/dark trials. However, despite the somewhat lengthened overall response latencies, the expected effect of word frequency was obtained during fMRI scanning in the current experiment, providing the opportunity to address the hypotheses discussed above.

Imaging Findings

As hypothesized, the parametric analyses revealed several areas of activation for word frequency, the three largest of which were left-lateralized (Figure 3.1): OT ($-47, -71, -3$), pSTG ($-51, -37, 20$), and anterior-inferior IFG ($-43, 29, 0$). Conjunction analysis revealed areas where the WFE showed a significant conjunction of frequency and familiarity in the following left-lateralized areas: posterior IT ($-44, -66, -4$), anterior insula ($-31, 24, 3$), and supramarginal gyrus (BA 40) (Figure 3.2). Overlapping of frequency and length effects was found exclusively in bilateral precentral gyrus ($-45, -13, 34; 49, -7, 36$) (Figure 3.2). Coordinates of activation for concept familiarity, but not frequency or word length, are reported in Table 1.

Table 1. Results from the Parametric Analysis of Concept Familiarity after Accounting for the Significant Effects of Word Frequency

Location	Size (Voxels)	Talairach Coordinates		
		x	y	z
<i>Temporal Lobe</i>				
Left OT	1439	-43	-60	-7
Right OT	189	47	-54	-9
Left paHG	70	-33	-40	-15
<i>Frontal Lobe</i>				
Left SFG	4339	-1	23	38
Left ant Ins	3837	-34	21	5
Left precent s	438	-46	4	31
Right ant Ins	65	30	22	1
<i>Parietal Lobe</i>				
Left SPL/IPS	320	-44	-43	48
Right SPL/IPS	29	42	-39	57
Right p cor rad	24	29	-57	22
Right SPL/IPS	15	25	-59	44
<i>Subcortical Nuclei</i>				
Left caudate	330	-8	4	2
Right thalamus	41	2	-23	11

OT = occipito-temporal cortex; STG = superior temporal gyrus; IFG = inferior frontal gyrus; SMG = supramarginal gyrus; SPL = superior parietal lobule; Ins = insula; IPL = inferior parietal lobule; paHG = parahippocampal gyrus; SFG = superior frontal gyrus; precent s = precentral sulcus; SPL/IPS = area of the superior parietal lobule/intraparietal sulcus; p cor rad = posterior region of the corona radiata.

A complete listing of Talairach coordinates for the centers of mass of the results of the parametric and conjunction analyses, including those in the parietal lobe not anticipated in our hypotheses, is provided in Tables 2 and 3, respectively.

Graphs of t -statistic values for the largest area showing a unique effect of word frequency, the pSTG, along with those for the three largest areas showing overlap of word frequency with at least one of the other two factors of interest, are provided in Figure 4. The colors in this graph correspond to those used to identify the effects of individual factors in Figure 3.

DISCUSSION

The cognitive model pictured in Figure 1 holds that the WFE primarily reflects access to lexical phonology,

Table 2. Results from the Parametric Analysis of Word Frequency after Accounting for the Significant Effects of Concept Familiarity and Word Length: *t*-Score, Size (in Voxels), and Talairach Coordinates (Talairach & Tournoux, 1988) for All Centers of Significant Activation ($p < .05$, Corrected)

Location	Size (Voxels)	Talairach Coordinates		
		<i>x</i>	<i>y</i>	<i>z</i>
<i>Temporal Lobe</i>				
Left OT	845	-47	-71	-3
Left posterior STG	448	-51	-37	20
<i>Frontal Lobe</i>				
Left IFG	192	-43	29	0
Right IFG	85	46	18	-7
Left precentral gyrus	18	-53	-7	34
<i>Parietal Lobe</i>				
Left SMG	121	-41	-39	45
Depth of right IPL	109	23	-61	36
Left SPL	86	-15	-70	50
Left SPL	66	-33	-53	53
Right SPL	52	19	-74	45
Right SPL	31	16	-68	50
Depth of left IPL	21	-33	-37	43
Depth of left IPL	19	-26	-54	45
<i>Cerebellum</i>				
Left medial cerebellum	32	-4	-75	-34

OT = occipito-temporal cortex; STG = superior temporal gyrus; IFG = inferior frontal gyrus; SMG = supramarginal gyrus; SPL = superior parietal lobule; IPL = inferior parietal lobule.

explaining the longer access times for lower-frequency words in terms of increased time needed for accessing lexical phonology. However, the specificity of the WFE is not entirely without controversy, particularly when there can be interaction of information between levels of representation. Additionally, for our set of stimuli, as is the case with most sets of pictorial stimuli, frequency is significantly correlated with both concept familiarity and word length. To address these concerns, we performed parametric analyses of concept familiarity and word length in order to separate lexical phonological effects from those of lexical semantics and articulatory-motor processing, respectively. The parametric analyses revealed a single area in the left pSTG that exhibited increased activation with production of increasingly low-frequency words and was not modulated by decreases

Table 3. Conjunction Analysis Results for Significant Overlap of Word Frequency and Concept Familiarity Effects (Top Rows), as well as Significant Overlap of Word Frequency and Word Length (Number of Syllables) Effects (Bottom Rows)

Location	Size (Voxels)	Talairach Coordinates		
		<i>x</i>	<i>y</i>	<i>z</i>
<i>Conjunction of Frequency and Familiarity Effects</i>				
Temporal lobe				
Left OT	704	-44	-66	-4
Frontal lobe				
Left anterior insula	474	-31	24	3
Parietal lobe				
Left SMG	11	-42	-41	47
<i>Conjunction of Frequency and Length Effects</i>				
Frontal lobe				
Left precentral gyrus	247	-45	-13	34
Right precentral gyrus	64	49	-7	36

OT = occipito-temporal cortex; SMG = supramarginal gyrus.

in familiarity or increases in word length (Table 2, Figures 3.1 and 4). This area was among the three largest showing activation for lower-frequency words. The other two, the left IFG and the OT, exhibited WFE-related activity which partially overlapped with concept familiarity effects (45.5% overlap for the left OT, 71.2% overlap for the left IFG/anterior insula). Completely overlapping frequency and length effects were observed bilaterally in the precentral gyri (Figure 3.2; slice $z = 36$). Centers of mass for the overlaps showed comparable levels of activation in the OT and IFG/anterior insula for word frequency and concept familiarity, and comparable word length and word frequency-related activity in a relatively large area of the left precentral gyrus (Figure 4).

Two prominent possible interpretations of these overlaps are that they represent a neural interaction of WFEs with those of related variables, or they could be due to a simple statistical confound (given that word frequency is significantly correlated with concept familiarity and word length in this dataset) that may not correspond to a meaningful neural relationship. Considering the case of distinguishing effects of word frequency from those of concept familiarity, if the overlap was due only to confounding of factors, we would expect that all regions showing a frequency effect would also show a familiarity effect, and vice versa. Instead, the data show, in addition to areas of overlap, several areas which appear to be modulated by either one factor or the other alone (e.g., pSTG for frequency, precentral sulcus for familiarity). Moreover, the behavioral analysis showed that familiarity

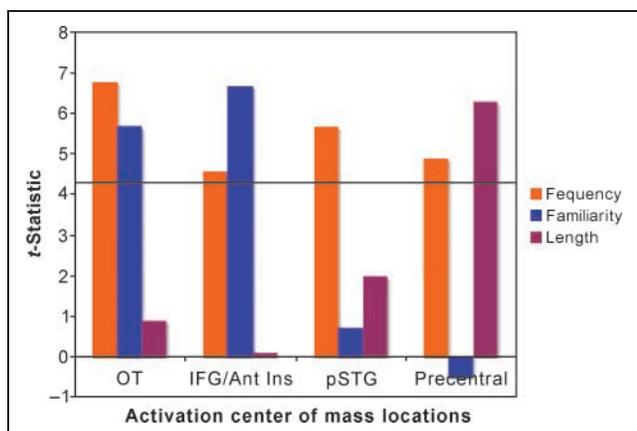


Figure 4. *t*-Statistic values for the respective centers of mass for areas showing a significant effect of word frequency alone (pSTG: $-51, -37, 20$, from Table 2), overlap of frequency and familiarity (OT: $-44, -66, -4$; IFG/anterior insula: $-31, 24, 3$, from Table 3), and overlap of frequency and length (precentral: $-45, -13, 34$, largest of the two precentral gyrus overlaps from Table 3). Horizontal line at approximately $t = 4.2$ corresponds to a significance threshold of $p < .05$, corrected.

and frequency both explained significant and unique proportions of latency variance. We infer that these factors tap into separable effects, and that the images of overlapping effects likely reflect true functional overlap.

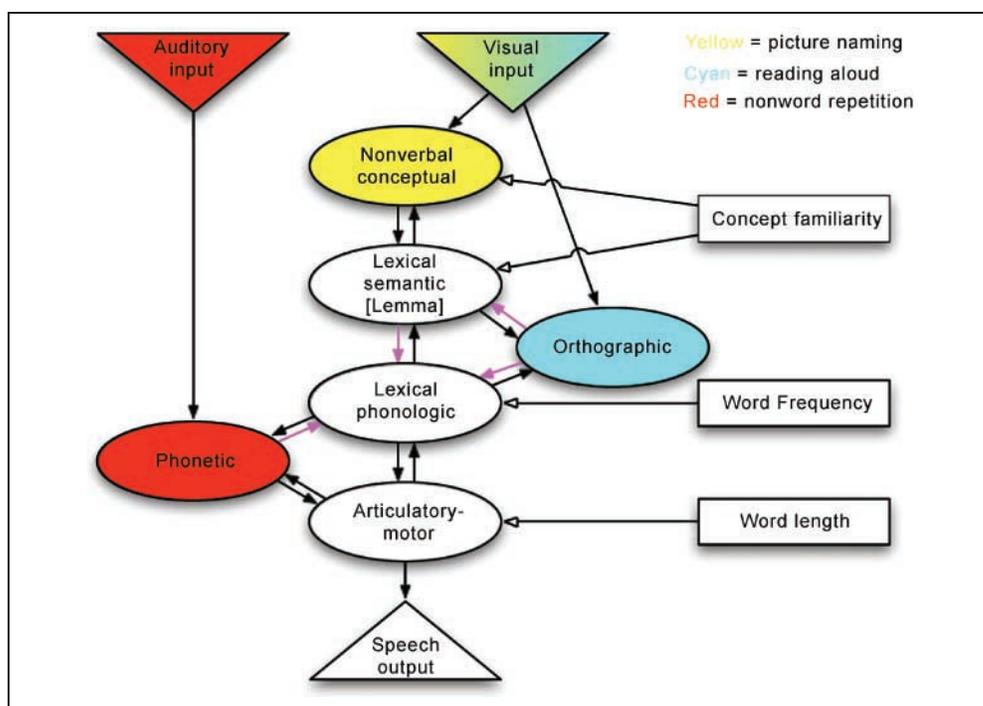
Temporal Lobe

The interpretation that the left pSTG is associated with access to the lexical phonological stage of word pro-

duction accords with two large-scale meta-analyses by Indefrey and Levelt (2000, 2004). Their evidence-based cognitive model places the WFE at the level of lexical phonological access, and their meta-analyses strongly implicate the left pSTG in this process. The current study confirms the prediction that the left pSTG supports access to lexical phonology by showing left pSTG activation specifically for the overt production of increasingly low-frequency words.

To our knowledge, the only other functional neuroimaging investigation of the WFE in overt word production is a PET study conducted by Fiez, Balota, Raichle, and Petersen (1999) using word stimuli instead of pictures. They also found increased activity in the left pSTG for low-frequency words. This convergence of results is accounted for in our extended word production model (Figure 5). According to this model and others (e.g., Indefrey & Levelt, 2000), picture naming (yellow in Figure 5) necessitates processing through the level of lexical semantics, whereas in word pronunciation (cyan in Figure 5), lexical semantic processing is optional (for a lucid review of WFEs in word pronunciation, please see Monsell, 1991). Thus, a conjunction analysis of WFEs across these two tasks, as is planned for future studies, should highlight areas more involved in lexical phonological than lexical semantic processing. Considering the Fiez et al. (1999) finding, along with that from the current study, the left pSTG appears to be consistently implicated in production of low-frequency words. More generally, Wise et al. (2001) analyzed a series of PET studies contrasting sound perception, word perception, and overt speech production. One

Figure 5. Version of the model shown in Figure 1 for picture naming, extended to account for processes thought to be involved in reading aloud and nonword repetition. As in Figure 1, ovals represent major representational stages, rectangles with open arrows represent parametric factors thought to preferentially modulate information at the indicated stages, and triangles represent input/output modalities. Magenta arrows represent levels at which word frequency is hypothesized to modulate information flow for the various tasks. Note that this model predicts that a conjunction across tasks would converge on a lexical phonologic level for word frequency effects (please see Discussion section for additional details).



area in which activations for these tasks converged was in the left pSTG, which they interpreted as activity at the interface between the recall and production of words. Thus, on the basis of the current study and others, it seems that the left pSTG is involved in generating sound patterns for words.

Converging evidence for the role of pSTG in the WFE comes from literature on repetition priming. Specifically, the WFE can be seen as a long-term priming effect, in which more frequently encountered words end up being primed more, and consequently, responded to more quickly, than less frequently encountered ones. This effect was demonstrated with pronounceable nonwords (pseudowords) by Majerus et al. (2005), who found increased activity in the left pSTG for production of less frequently presented pseudowords compared to those which were more frequently encountered. Because pseudowords, by definition, do not refer to meaningful entries in the lexicon, such differential activity would presumably not arise from lexical semantic level processes.

In addition to temporal lobe findings in the pSTG, we also found an association for production of low-frequency words with activation in the left OT (OT, labeled as activation A in Figure 3, which lies in the posterior IT, roughly lateral BA 37/19). Just ventral and anterior to this is a region roughly corresponding to the fusiform gyrus (BA 37 and posterior 20), which has been consistently implicated in the retrieval of either conceptual or lexical semantic information in visual and verbal tasks (Martin & Chao, 2001; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996), including picture naming (Murtha et al., 1999). For example, the study by Vandenberghe et al. (1996), in which subjects made semantic association judgments on both pictures and words, showed activation in the left fusiform gyrus. It is possible that a posterior-to-anterior gradient exists in which aspects of phonological processing occur in the posterior OT, semantic processing in the fusiform gyrus, and a co-occurrence of the two, possibly to facilitate binding, in the anterior OT. Such binding may be reflected by the overlapping activation for word frequency and concept familiarity factors shown in Figures 3.2 and 4.

There is also evidence of a more phonology-specific role for the OT. In a pseudoword pronunciation paradigm in which subjects were trained to recognize a set of pseudowords using either phonological, semantic, or orthographic information, subjects exhibited greatest performance (shorter latencies to pronunciation) following training in the phonological condition. Reduced neural activity was also observed in the left OT during word pronunciation following phonological training, when compared to word pronunciation following semantic or orthographic training (Sandak et al., 2004). However, findings from the word pronunciation task used by Sandak et al. may not generalize to the current picture naming study. The confound of material type seems particularly problematic for interpreting activa-

tion in the OT, which has been dubbed the “visual word form area” (McCandliss, Cohen, & Dehaene, 2003; Cohen et al., 2002). In general, more study would seem to be warranted before any strong conclusions can be made regarding the specifics of the role of the OT in word pronunciation.

The temporal lobe activations seen here in the pSTG and OT associated with production of increasingly low-frequency words appear to be playing somewhat different roles. We hypothesize that the neural activity in the OT for the WFE and its overlap with familiarity effects reflects the retrieval of heteromodal, intermediary-stage lexical representations. This information would be heteromodal in the sense that it contains semantic information derived from the binding of elements from multiple sensory modalities, and lexical in the sense that it has been abstracted to the point where it can be paired with a particular phonological word form. Such information would essentially be the equivalent of what Levelt and colleagues have termed the lemma. Activity in the pSTG, on the other hand, is proposed to reflect activation of the corresponding phonological word form.

Frontal Lobe

Using a noninteractive version of a model similar to that shown in Figure 1, Indefrey and Levelt (2004) analyzed a combination of chronometric and functional imaging studies of word production. They concluded that activation in the left IFG could begin during a time window overlapping or closely following that for the left pSTG, and most likely corresponded to the syllabification stage in the production of a word’s phonology. This is consistent with interpreting activation in the anterior left IFG (roughly BA 47/45) found in the current study for production of increasingly low-frequency words as being involved in aspects of lexical phonological access.

However, our finding of activation in BA 47/45 is anterior and ventral to what would be predicted from models which propose domain-specific processing within the left IFG (Bookheimer, 2002; Poldrack et al., 1999; Fiez, 1997); that is, prediction 1 (see Introduction) was not borne out. Put another way, domain-specific accounts would hold that the activation observed in the anterior–ventral IFG for a language task is related to semantic access rather than to lexical phonological access. Prediction 4—that word frequency and concept familiarity effects would interact in the anterior–ventral IFG, consistent with a domain-general view—was borne out. Note, however, that neither the domain-specific nor the domain-general view would seem to predict WFE-specific (i.e., lexical phonological) activation in the anterior–ventral IFG in the absence of related activation in the posterior–dorsal IFG. Also, these accounts deal specifically with the IFG, whereas our results show an overlap of word frequency and concept familiarity effects in the anterior insula adjacent to the WFE-specific

activation in the IFG. We hypothesize that the left-sided IFG, OT, and possibly anterior insula, act together to make both lexical semantic and some aspects of lexical phonological information available during naming, with frontal areas directing the activity in the temporal lobe for retrieving heteromodal, intermediary representations such as the lemma.

Parietal Lobe

In addition to the areas discussed above, production of low-frequency words was also associated with bilateral activations in the area of the intraparietal sulcus/superior parietal lobule, in addition to a single WFE-specific activation in the left supramarginal gyrus (SMG) and an area of overlap for word frequency and concept familiarity effects in a smaller part of the same area. This finding is in contrast to the consistent lack of parietal activation seen in the large meta-analysis of word production conducted by Indefrey and Levelt (2004).

Because parietal activations were not specifically hypothesized, it is difficult to make strong conclusions from results of the current experiment alone. However, given that activity in the parietal lobe has been previously associated with eye movements in visual search (Ipata, Gee, Goldberg, & Bisley, 2006), the 1000-msec presentation duration used in this study might have allowed enough time for differences in exploratory eye movements to affect the results.

Future studies may gain insight into the possible cognitive implications of patterns of parietal lobe activations such as those seen here through use of the extended model shown in Figure 5. This model provides a framework for interpreting WFEs across tasks, where converging results would provide strong evidence for activation specifically reflecting access to lexical phonology. This model is not all-encompassing, however, and factors which are beyond its scope might also elicit activation such as that seen here in the parietal lobes. Attention, a factor which has also been shown to modulate parietal activity (Serences & Yantis, 2006), could, in principle, correlate with increased demands at some of the processing levels shown in Figure 5, as could additional external factors such as eye movements.

Conclusion

Using the WFE as a probe for lexical phonological level processing, along with concept familiarity and word length effects to probe for lexical semantic and articulatory-motor processing, respectively, we provide evidence that the left pSTG is specifically involved in accessing lexical phonology. We also found a distinct set of areas, the left IFG and the OT, to be related to both lexical phonology and lexical semantics. Based on these results, we propose a model for overt picture naming in which the pSTG acts as a convergence region to coordinate

and hold on line the lexical phonological form corresponding to semantic information recalled through activity in the OT and, possibly, the IFG. Once activated, this lexical phonological form can be paired with associated information such as its syllabic content and articulatory score, perhaps through the coordinated activity of the IFG and the precentral gyri. The specifics of whether areas showing an overlap of frequency and familiarity effects reflect the activation of intermediary lemma-level information, as well as the exact nature of any interaction these overlaps might represent, will be the subject of future research.

Acknowledgments

We thank Hanna Damasio, Prahlad Gupta, Daniel Tranel, Steven W. Anderson, David Rudrauf, and Antonio R. Damasio for extremely valuable discussion and willingness to share ideas. Additionally, we thank David Rudrauf for technical assistance with Figure 3. Support for this research was contributed by NIH grants from the NIDCD to T. J. G. (R01 DC006633 and R33 EB001484) and an NRSA to W. W. G. (F31 DC007560).

Reprint requests should be sent to William W. Graves, Medical College of Wisconsin, Language Imaging Lab, MEB 4550, 8701 Watertown Plank Rd., Milwaukee, WI 53226, or via e-mail: wgraves@mcw.edu.

REFERENCES

- Alario, F., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140–155.
- Alexander, M. P. (2003). Aphasia: Clinical and anatomic issues. In T. E. Feinberg & M. J. Farah (Eds.), *Behavioral neurology and neuropsychology* (2nd ed., pp. 147–164). New York: McGraw-Hill.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. 2.5.
- Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews*, *1*, 206–218.
- Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., et al. (2003). Neural correlates of lexical access during visual word recognition. *Journal of Cognitive Neuroscience*, *15*, 372–393.
- Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151–188.
- Broca, P. (2006). Comments regarding the seat of the faculty of spoken language, followed by an observation of aphemia (loss of speech) (S. Karir, Trans.). In Y. Grodzinsky & K. Amunts (Eds.), *Broca's region* (pp. 291–304). New York: Oxford University Press.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*, 177–208.
- Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, *57*, 8–16.
- Chee, M. W. L., Hon, N. H. H., Caplan, D., Lee, H. L., & Goh, J. (2002). Frequency of concrete words modulates prefrontal activation during semantic judgments. *Neuroimage*, *16*, 259–268.

- Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, *125*, 1054–1069.
- Coughlan, A. K., & Warrington, E. K. (1978). Word-comprehension and word-retrieval in patients with localized cerebral lesions. *Brain*, *101*, 163–185.
- Damasio, A., & Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In C. Koch & J. Davis (Eds.), *Large-scale neuronal theories of the brain* (pp. 61–74). Cambridge: MIT Press.
- Damasio, A. R. (1989a). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, *1*, 123–132.
- Damasio, A. R. (1989b). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25–62.
- Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D., & Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature*, *380*, 499–505.
- Damasio, H., Tranel, D., Grabowski, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*, 179–229.
- Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes*, *5*, 313–349.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*, 801–838.
- Fiebach, C. J., Friederici, A. D., Muller, K., & Yves von Cramon, D. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, *14*, 11–23.
- Fiez, J. A. (1997). Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human Brain Mapping*, *5*, 79–83.
- Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. *Neuron*, *24*, 205–218.
- Fiez, J. A., & Tranel, D. (1997). Standardized stimuli and procedures for investigating the retrieval of lexical and conceptual knowledge for actions. *Memory & Cognition*, *25*, 543–569.
- Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*, 636–647.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*, 627–635.
- Frank, R. J., Damasio, H., & Grabowski, T. J. (1997). Brainvox: An interactive, multimodal visualization and analysis system for neuroanatomical imaging. *Neuroimage*, *5*, 13–30.
- Gold, B. T., & Buckner, R. L. (2002). Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. *Neuron*, *35*, 803–812.
- Gordon, J. K. (2002). Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language*, *82*, 113–145.
- Grabowski, T. J., Bauer, M. D., Foreman, D., Mehta, S., Eaton, B. L., Graves, W. W., et al. (2006). Adaptive pacing of visual stimulation for fMRI studies involving overt speech. *Neuroimage*, *29*, 1023–1030.
- Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*, 313–338.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, *5*, 67–103.
- Indefrey, P., & Levelt, W. J. M. (2000). The neural correlates of language production. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 845–865). Cambridge: MIT Press.
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.
- Ipata, A. E., Gee, A. L., Goldberg, M. E., & Bisley, J. W. (2006). Activity in the lateral intraparietal area predicts the goal and latency of saccades in a free-viewing visual search task. *Journal of Neuroscience*, *26*, 3656–3661.
- Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 824–843.
- Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 432–438.
- Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *Neuroimage*, *21*, 946–953.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 83–122). New York: Oxford University Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–75.
- Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*, 122–142.
- Majerus, S., Van der Linden, M., Collette, F., Laureys, S., Poncelet, M., Degueldre, C., et al. (2005). Modulation of brain activity during phonological familiarization. *Brain and Language*, *92*, 320–331.
- Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology*, *11*, 194–201.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences*, *7*, 293–299.
- McDermott, K. B., Petersen, S. E., Watson, J. M., & Ojemann, J. G. (2003). A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. *Neuropsychologia*, *41*, 293–303.
- Mechelli, A., Crinion, J. T., Long, S., Friston, K. J., Lambon Ralph, M. A., Patterson, K., et al. (2005). Dissociating

- reading processes on the basis of neuronal interactions. *Journal of Cognitive Neuroscience*, *17*, 1753–1765.
- Mehta, S., Grabowski, T. J., Razavi, M., Eaton, B., & Bolinger, L. (2006). Analysis of speech-related variance in rapid event-related fMRI using a time-aware acquisition system. *Neuroimage*, *29*, 1278–1293.
- Mesulam, M.-M. (1998). From sensation to cognition. *Brain*, *121*, 1013–1052.
- Monsell, S. (1990). Frequency effects in lexical tasks: Reply to Balota and Chumbley. *Journal of Experimental Psychology: General*, *119*, 335–339.
- Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 148–197). Hillsdale, NJ: Erlbaum.
- Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, *20*, 705–714.
- Moss, H. E., Abdallah, S., Fletcher, P., Bright, P., Pilgrim, L., Acres, K., et al. (2005). Selecting among competing alternatives: Selection and retrieval in the left inferior frontal gyrus. *Cerebral Cortex*, *15*, 1723–1735.
- Murtha, S., Chertkow, H., Beaugard, M., & Evans, A. (1999). The neural substrate of picture naming. *Journal of Cognitive Neuroscience*, *11*, 399–423.
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*, 653–660.
- Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*, 273–281.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*, 585–589.
- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 765–805.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*, *10*, 15–35.
- Rapp, B., & Goldrick, M. (2006). Speaking words: Contributions of cognitive neuropsychological research. *Cognitive Neuropsychology*, *23*, 39–73.
- Raymer, A. M., Foundas, A. L., Maher, L. M., Greenwald, M. L., Morris, M., Rothi, L. J. G., et al. (1997). Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia. *Brain and Language*, *58*, 137–156.
- Saffran, E. M. (2000). Aphasia and the relationship of language and brain. *Seminars in Neurology*, *20*, 409–418.
- Sandak, R., Mencl, W. E., Frost, S. J., Rueckl, J. G., Katz, L., Moore, D. L., et al. (2004). The neurobiology of adaptive learning in reading: A contrast of different training conditions. *Cognitive, Affective, & Behavioral Neuroscience*, *4*, 67–88.
- Schriefers, H., Meyer, A. S., & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language*, *29*, 86–102.
- Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, *10*, 38–44.
- Shattuck-Hufnagel, S. (1992). The role of word structure in segmental serial ordering. *Cognition*, *42*, 213–259.
- Smyser, C., Grabowski, T. J., Frank, R. J., Haller, J. W., & Bolinger, L. (2001). Real-time multiple linear regression for fMRI supported by time-aware acquisition and processing. *Magnetic Resonance in Medicine*, *45*, 289–298.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 174–215.
- Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, *28*, 516–536.
- Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, *14*, 17–26.
- Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *CRL Newsletter*, *12*, 3–33.
- Székely, A., D'amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., et al. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, *35*, 621–633.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic axis of the human brain*. Stuttgart: Thieme.
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring “how” from “where.” *Neuropsychologia*, *41*, 280–292.
- Thompson-Schill, S. L., Bedny, M., & Goldberg, R. F. (2005). The frontal lobes and the regulation of mental activity. *Current Opinion in Neurobiology*, *15*, 219–224.
- Tranel, D., Damasio, H., & Damasio, A. R. (1997). A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia*, *35*, 1319–1327.
- Tranel, D., Damasio, H., & Damasio, A. R. (1998). The neural basis of lexical retrieval. In R. W. Parks, D. S. Levine, & D. L. Long (Eds.), *Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience* (pp. 271–296). Cambridge: MIT Press.
- Tranel, D., Grabowski, T. J., Lyon, J., & Damasio, H. (2005). Naming the same entities from visual or from auditory stimulation engages similar regions of left inferotemporal cortices. *Journal of Cognitive Neuroscience*, *17*, 1293–1305.
- Vandenberghe, R., Price, C., Wise, R. J. S., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*, 254–256.
- Vitevitch, M. S. (2002). The influence of phonological similarity neighborhoods on speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 735–747.
- Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *American Journal of Psychology*, *81*, 226–234.
- Wise, R. J. S., Scott, S. K., Blank, C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within “Wernicke’s area.” *Brain*, *124*, 83–95.
- Woods, R. P., Dapretto, M., Sicotte, N. L., Toga, A. W., & Mazziotta, J. C. (1999). Creation and use of a Talairach-compatible atlas for accurate, automated, nonlinear intersubject registration, and analysis of functional imaging data. *Human Brain Mapping*, *8*, 73–79.
- Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., & Mazziotta, J. C. (1998). Automated image registration: I. General methods and intrasubject, intramodality validation. *Journal of Computer Assisted Tomography*, *22*, 139–152.
- Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., & Mazziotta, J. C. (1998). Automated image registration: II. Intersubject validation of linear and nonlinear models. *Journal of Computer Assisted Tomography*, *22*, 153–165.