

# Top-Down Influences on Lexical Selection During Spoken Word Production: A 4T fMRI Investigation of Refractory Effects in Picture Naming

Greig de Zubicaray,\* Katie McMahon, Mathew Eastburn, and Alan Pringle

*Centre for Magnetic Resonance, University of Queensland, Queensland, Australia*

**Abstract:** Spoken word production is assumed to involve stages of processing in which activation spreads through layers of units comprising lexical-conceptual knowledge and their corresponding phonological word forms. Using high-field (4T) functional magnetic resonance imaging (fMRI), we assessed whether the relationship between these stages is strictly serial or involves cascaded-interactive processing, and whether central (decision/control) processing mechanisms are involved in lexical selection. Participants performed the competitor priming paradigm in which distractor words, named from a definition and semantically related to a subsequently presented target picture, slow picture-naming latency compared to that with unrelated words. The paradigm intersperses two trials between the definition and the picture to be named, temporally separating activation in the word perception and production networks. Priming semantic competitors of target picture names significantly increased activation in the left posterior temporal cortex, and to a lesser extent the left middle temporal cortex, consistent with the predictions of cascaded-interactive models of lexical access. In addition, extensive activation was detected in the anterior cingulate and pars orbitalis of the inferior frontal gyrus. The findings indicate that lexical selection during competitor priming is biased by top-down mechanisms to reverse associations between primed distractor words and target pictures to select words that meet the current goal of speech. *Hum Brain Mapp* 27: 864–873, 2006. © 2006 Wiley-Liss, Inc.

**Key words:** fMRI; language production; semantic processing; inhibition; competitor priming

## INTRODUCTION

The selection of lexical information is a key requirement for producing speech. In most contemporary theories of spoken word production, access to this information occurs at two discernible processing stages. These stages comprise,

respectively, units of combined lexical-conceptual and syntactic knowledge (referred to as “lemmas” in some theories) and their accompanying phonological word forms [Dell and Sullivan, 2004; Levelt, 2001]. Although theoretical accounts tend to agree about the deployment of these two stages of lexical representations, proposals differ markedly with respect to the nature of the processing that occurs between them. Three possible mechanisms have been proposed and implemented in computational models: serial, cascaded and interactive processing.

Contract grant sponsor: Australian Research Council (ARC) Discovery Project; Contract grant number: DP0342945.

\*Correspondence to: Greig de Zubicaray, Centre for Magnetic Resonance, The University of Queensland, QLD, 4072, Australia. E-mail: greig.dezubicaray@cmr.uq.edu.au

Received for publication 19 July 2005; Revised 2 November 2005; Accepted 3 November 2005

DOI: 10.1002/hbm.20227

Published online 17 March 2006 in Wiley InterScience (www.interscience.wiley.com).

Serial models make the assumption that lexical selection occurs solely at the lexical-conceptual (lemma) stage, with processing continuing in a strictly feed-forward manner through remaining levels until articulation takes place [e.g., Levelt et al., 1999]. Cascade models alternatively assume that lexical-conceptual and word form information can co-activate as soon as possible, i.e., although lexical-conceptual

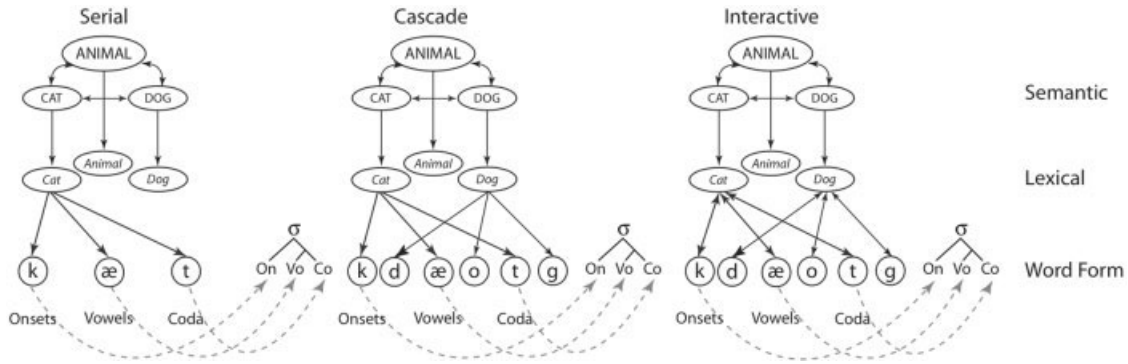


Figure 1.

Examples of serial [Levelt et al., 1999], cascaded [Peterson and Savoy, 1998], and interactive [Dell and Sullivan, 2004] network models consisting of semantic, lexical (lemma) and phonological units (word forms). In the serial model, a single target lemma (e.g., *cat*) is selected from among semantic competitors (e.g., *dog*) and

encoded phonologically, whereas in the cascade model activation spreads to all phonological units before lexical selection is completed. In the interactive model, activation also spreads from the phonological to the lexical units.

processing is the first to commence, there is a period during which all of these units are activated and their corresponding word forms retrieved, with the former activity ending before word form retrieval is concluded [Morsella and Miozzo, 2002; Peterson and Savoy, 1998]. In addition to assuming cascaded processing, models incorporating interactivity make the assumption that activated word forms can excite their lexical-conceptual units, usually via feedback. These models presuppose cascaded processing as phonological retrieval must occur in advance to influence lexical selection [Cutting and Ferreira, 1999; Dell and O’Seaghdha, 2001; Harley, 1993]. All may be considered connectionist “in the sense that computation is carried out by spreading activation through a network of units representing lexical knowledge” [Dell and Sullivan, 2004; p. 68]. Figure 1 illustrates the network components of the three different model types.

the lexical-conceptual level [e.g., Levelt et al., 1999], whereas interactive models assume this happens at word form retrieval [e.g., Harley, 1993, Starreveld and La Heij, 1996].

### Refractory Effects in Spoken Word Production: The Competitor Priming Paradigm

All of the above models assume competition occurs naturally between a target and its lexical neighbors during processing. This is supported by analyses of occasional errors in normal speech showing semantic substitutions, in which an intended word (e.g., *warm*) is replaced by a semantically related word [e.g., *cold*; Harley and MacAndrew, 2001]. Promoting competition between a target word and its lexical neighbors, particularly during picture-naming tasks, is an established method for investigating the processes involved in spoken word production. A prominent example is the picture-word interference paradigm, in which a distractor word is presented concurrently with a pictured target object [e.g., Damian and Martin, 1999]. Distractor words from the same semantic category as the target (e.g., picture DOG, word *cat*) result in slower naming latencies compared to an unrelated word (e.g., picture DOG, word *axe*). According to theoretical models, presentation of a semantically related distractor word further excites its already activated units such that lexical selection of the target is slowed until the stronger competition is resolved. Serial models, however, assume that this competition occurs and is resolved at

Semantic interference effects in word production can also be obtained with the competitor priming paradigm, introduced by Wheeldon and Monsell [1994]. The paradigm requires participants to generate a categorically related distractor word (the prime, e.g., *whale*) in response to a definition (e.g., “the largest creature that swims in the sea”) before their naming a picture of a target object (SHARK). Wheeldon and Monsell [1994] observed slower naming latencies when two trials intervened between the primed distractor word and the target picture. As per explanations of picture-word interference effects, the prime item is assumed to be a more potent competitor due to its activation level being increased such that lexical selection of the subsequently presented target is delayed [Belke et al., 2005; Tree and Hirsh, 2003; Wheeldon and Monsell, 1994]. An additional assumption of refractoriness of the prime’s activation level is required to account for it being maintained across the intervening trials [see Forde and Humphreys, 1997].

Wheeldon and Monsell [1994] suggested the lexical-conceptual level of processing to be the most likely candidate for the source of the competitor priming effect [see also Belke et al., 2005; Tree and Hirsh, 2003]. They emphasized, however, that they made no claim about the relationship between lemma selection and retrieval of the corresponding word form (i.e., serial, cascaded, or interactive processing mechanisms). In fact, some evidence suggests that the competitor priming effect might instead be due to competition at word form retrieval, consistent with the assumptions of interactive models of spoken word production [e.g., Harley, 1993; Starreveld and La

Heij, 1996]. Using Dutch participants, Wheeldon [2003] demonstrated that naming latencies to a target object were slowed after the generation of a phonologically related word in response to a definition on an earlier trial. Ferreira and Griffin [2003] showed that homophones of semantic competitors (e.g., *none*) primed by cloze sentences on previous trials elicited word substitution errors to target objects (e.g., *PRIEST*) to a similar extent as primed semantic competitors (e.g., *nun*). They concluded that this outcome was consistent with lexical-conceptual units and word forms being co-activated, with activation of the word forms influencing lemma selection, i.e., evidence of interactivity in spoken word production [see also Cutting and Ferreira, 1999].

### Neuroimaging Tests of Models of Lexical Access in Spoken Word Production

Indefrey and Levelt [2000, 2004] conducted a meta-analysis of functional neuroimaging studies of spoken word production, relating the data to the stages implemented in Levelt et al.'s [1999] serial model. The meta-analysis identified reliable roles for the midsection of the middle temporal gyrus in lexical-conceptual processing (lemma selection) and the posterior superior and middle temporal gyri (Wernicke's area) in word form retrieval, respectively, within a predominantly left-hemisphere-based cerebral network. Importantly, the results of this meta-analysis are generalizable to other (cascaded and interactive) models of spoken word production as differences in model assumptions relate only to the nature of the processes occurring between stages, not to the stages themselves [Indefrey and Levelt, 2004].

Following from our investigations of picture-word interference effects with functional magnetic resonance imaging (fMRI) that provided support for multiple stages of lexical access [see Dell and Sullivan, 2004], we offered a revised account of lexical access in spoken word production in which Botwinick et al.'s [2001] conflict monitoring and cognitive control layers were appended to Harley's [1993] interactive model [de Zubicaray, in press; de Zubicaray et al., 2001, 2002]. Conflict is defined operationally therein as the simultaneous activation of incompatible, mutually inhibiting representations. Once conflict is detected, the control unit intervenes to guide the activation in the network to meet the task goals (in this case, naming a target object). Botwinick et al. [2001] had added these layers successfully to several well-known examples of interactive models of information processing, including the Stroop [1935] task. They also identified conflict monitoring as one function of the anterior cingulate cortex [ACC; see Barch et al., 2001; Botwinick et al., 2004].

As they often incorporate inhibitory links within or between the stages of lexical representation, interactive models of spoken word production provide a mechanism for suppressing a distractor word's activated lexical units, preventing their selection [see Berg and Schade, 1992b; Harley, 1993; Miozzo and Caramazza, 2003]. The use of these links for competitor deactivation has been a contentious issue [Berg and Schade, 1992a; Dell and O'Seaghdha,

1994; Miozzo and Caramazza, 2003], as other models assume only that the activation levels of competitors return to a resting state via a decay-based mechanism once the lexical unit with the highest level of activation has been selected [e.g., Levelt et al., 1999]. Ferreira and Pashler's [2002] finding of central bottlenecks during lemma selection and word form retrieval indicates that decision or control processes operate at both stages of lexical access. As per our account of picture-word interference, we assume control processes intervene in the competitor priming paradigm by manipulating the activation levels of strong lexical competitors through the use of inhibitory links to reverse an association developed between the primed distractor word and the target object's name.

Inhibitory control has long been considered a function of the prefrontal cortex (PFC); however, there is disagreement regarding the precise subregion that might be responsible for this function. Two prominent views in the cognitive neuroscience literature attribute inhibitory control to the inferior frontal cortex (in particular, the pars opercularis in the inferior frontal gyrus; IFG [Aron et al., 2004]) and orbitofrontal cortex [OFC; Fuster, 1997, 2005], respectively. Although these regions have been implicated across a diverse range of cognitive and motor tasks, the evidence available in terms of inhibitory control in spoken word production is scant. Duffau et al. [2005], however, demonstrated recently that electrical stimulation/interruption of the pars orbitalis of the IFG (considered part of the OFC [Ridderinkhof et al., 2004]) during intraoperative language mapping elicited semantic substitutions during picture naming. Moreover, semantic substitutions were also elicited by stimulation of the posterior superior and middle temporal gyri, the regions identified by Indefrey and Levelt's [2004] meta-analysis as being involved in word form retrieval.

We conducted an fMRI experiment using Wheeldon and Monsell's [1994] competitor priming paradigm. A compressed or sparse temporal sampling design was employed that permitted participants to make overt naming responses to target pictures [see de Zubicaray et al., 2001, 2002]. This design can detect changes in cerebral activity in a robust manner comparable to continuous imaging [Nebel et al., 2005]. It has the distinct advantage, however, of permitting the utterance and recording of spoken words without generating speech-related motion-induced signal artifact [Gracco et al., 2005]. We tested three predictions following from our interactive account of competitive priming described above. First, interactive models assume that refractory effects in picture naming occur due to competition during retrieval of activated word forms [e.g., Harley, 1993], whereas serial models assume this is due to competition during lexical-conceptual processing [lemma selection; Levelt et al., 1999]. According to the results of Indefrey and Levelt's [2004] meta-analysis, this should be revealed as increased activation in either the left posterior superior and middle temporal gyri, if the former account is correct, or in the midsection of the left middle temporal gyrus. A conflict-monitoring role has been associated with the functions of the ACC [Botwinick et al., 2001, 2004]. Hence, if conflict monitoring is

involved in competitor priming due to the activation of mutually inhibiting lexical representations, this should be revealed as an increase in ACC activation. Finally, if inhibitory control is engaged once conflict is detected to manipulate the activity levels of lexical competitors, then this should be reflected in increased activation in either the pars opercularis or pars orbitalis, although we favor the latter region given Duffau et al.'s [2005] results mentioned above.

## SUBJECTS AND METHODS

### Participants

Thirteen healthy subjects (seven males) with a mean age of 26.8 years (standard deviation [SD] 4.7) participated in the experiments. They were recruited from undergraduate and postgraduate students of the University of Queensland. All were right-handed native English speakers, with no history of neurological or psychiatric disorder or substance dependence, and all had normal hearing and vision. They provided written informed consent in accordance with procedures approved by the University of Queensland's Medical Research Ethics Committee, and were reimbursed AUD \$30 for their participation.

### Materials

The materials consisted of 50 pairs of semantically related nouns taken from Wheeldon and Monsell [1994] matched according to a range of variables (see their Appendix 1). An experimental list was constructed that consisted of 25 of these pairs from a number of semantic categories (e.g., animals, musical instruments, household items, etc). One word in each pair was elicited by a picture (the target object name) and the other by a definition (the prime word). The pictures were black-and-white line drawings taken from several sources, including Snodgrass and Vanderwart [1980], Szekely et al. [2004], and Bonin et al. [2003]. For each experimental pair, two additional words were selected, unrelated to the target and from a different semantic category to be unprimed controls [or "fillers"; Wheeldon and Monsell, 1994] elicited on the two trials between prime and target: one by a picture of an object, the other by a definition. The experimental and control groups of word pairs were further matched (all  $F_s < 1$ ) according to word length, number of phonemes, number of syllables, and mean frequencies of occurrence per million in the CELEX English Database [1993] (Table I). A further pair of words to be elicited by an unrelated definition and a picture was selected for a practice trial before the experimental session.

### Apparatus

A PC running Microsoft *VisualBasic* and *Exacticks* (Ryle Design) software was used to deliver the picture stimuli and definitions and record vocal responses on digital audio files (sampling rate 11 kHz). Text and line drawings were presented in black on a luminous white background, enlarged

**TABLE I. Properties of the matched definition and target word pairs for primed and unprimed conditions**

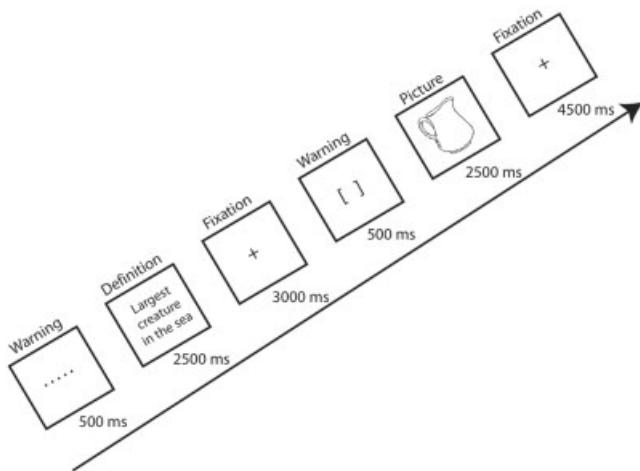
Properties	Primed		Unprimed	
	Definition	Target	Definition	Target
Mean phonemes	3.52	3.92	3.52	3.96
Mean syllables	1.36	1.48	1.24	1.44
Mean frequency	52.08	42.28	49.24	40.32
Mean letters	4.92	5.00	4.44	4.92

and back-projected onto a screen that the participants viewed through a mirror mounted on the head coil using a BenQ SL705X projector. The stimuli subtended approximately 10 degrees of visual arc when each participant was positioned for imaging. Naming responses were recorded using a custom-positioned magnetic resonance (MR)-compatible microphone attached to the head coil. Naming latencies were then determined with conventional voice-key software. The filtered audio files were consulted for scoring of each participant's responses.

### Procedure

Participants were familiarized with the set of experimental and control pictures by viewing each one on a computer screen in random order with the appropriate label printed below. The size of the pictures including background was approximately 10 cm wide by 10 cm high. This occurred over two consecutive practice blocks in which they were instructed to name the pictures as fast and as accurately as possible. Erroneous naming responses were corrected. In a third practice block, they viewed the pictures without the labels printed below and were instructed to name the pictures as per the previous instructions. After this, participants were provided with examples of the definition stimuli to be used. Before the experiment, they were instructed to provide single word responses as quickly and accurately as possible to the definitions and pictures. In the event of an error, they were instructed not to correct their response and to wait for the next trial. Participants were also instructed not to speak or move during image acquisition (as indicated by the relatively loud gradient noise).

The experimental session consisting of a single block of 50 trials, preceded by one practice trial, was then conducted. This consisted of 25 target pictures in each of two conditions: unprimed (control) and primed by a response to a definition after two unrelated trials (one picture and one definition, referred to as "lag = 2" by Wheeldon and Monsell [1994]). The presentation of a trial involved the following sequence: Participants were presented with a warning display consisting of a row of dots (...) for 500 msec. This indicated a definition was about to appear. At the offset of the warning, a definition was displayed for 2,500 msec allowing the participant to respond, followed by a fixation point (+) for 3,000 msec. Another warning display was then presented, consisting of a pair of closed square brackets ([ ]) indicating that a picture was about to be displayed. At the offset of this



**Figure 2.**

A typical sequence involving presentation of a definition and picture in the competitor priming task. A single image volume (repetition time 3 seconds) was acquired after presentation of the picture and the participant’s naming response.

warning, the picture was displayed for 2,500 msec to allow the participant to respond, followed by a fixation point for 4500 ms. During the last 3000 msec of this fixation period, a single image volume was acquired; thus, the time between the response with the prime word and presentation of the target picture was approximately 19.5 sec. A schematic of the experimental design is shown in Figure 2.

### Image Acquisition

Participants were imaged with a Bruker Medspec system operating at 4 Tesla using a transverse electromagnetic (TEM) head coil for radiofrequency transmission and reception [Vaughn et al., 2002]. A point-spread function (PSF) mapping sequence was acquired before the echo planar imaging (EPI) time-series acquisitions to correct geometric distortions in the high-field EPI data [Zaitsev et al., 2003; Zheng and Constable, 2002]. A gradient echo EPI sequence optimized for both image quality and noise reduction [McMahon et al., 2004] was then used to acquire T2\*-weighted images depicting blood oxygenation level-dependent (BOLD) contrast ( $64 \times 64$  matrix;  $3.6 \times 3.6$  mm voxels). In a single session, 52 image volumes of 36 axial 3.5-mm slices (0.1 mm gap) were acquired (effective repetition time, 13.5 sec; echo time, 30 msec; flip angle, 60 degrees). The first two volumes were discarded. Behavioral trials were interleaved with image acquisition using sparse temporal sampling [Gracco et al., 2005] to capture the estimated peak BOLD signal response to task-related neural activity (time-to-peak approximately  $4.7 \pm 1.1$  sec [Aguirre et al., 1998]). For each trial, no field gradients were applied for a 11.5-sec period of silence allowing for stimuli to be presented and the participant’s overt verbal response, then immediately applied for image acquisition. A single image volume was then acquired

within 3 sec, approximately coincident with the picture naming trial’s estimated peak BOLD response. Total imaging time was approximately 25 min. Head movement was limited by foam padding within the head coil. Finally, a high-resolution 3D T1-weighted image was acquired using a magnetization prepared rapid acquisition gradient echo sequence (MPRAGE;  $256^3$  matrix;  $0.9 \text{ mm}^3$  voxels).

### Image Processing and Analysis

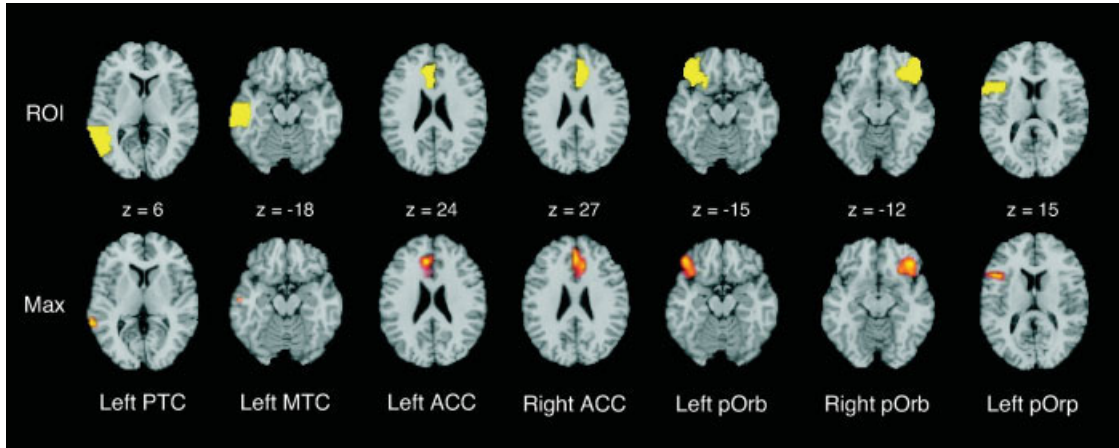
Rigid body motion correction of the fMRI time-series data was carried out using *INRIAlign*, and a mean image was generated from the realigned series [Freire et al., 2002]. The realigned data was then regrouped so that the images from each condition were treated as a single epoch and trials meeting exclusion criteria removed (see below). Further image preprocessing and statistical analyses were carried out using statistical parametric mapping software (*SPM2*; Wellcome Department of Cognitive Neurology, Queens Square, London, UK). The mean image from each participant was spatially normalized via nonlinear basis functions to the corresponding *SPM2* EPI template image in the Montreal Neurological Institute (MNI) atlas space [Ashburner and Friston, 1999; Evans et al., 1994]. The nonlinear transformations were next applied to the realigned and regrouped time-series volumes from which the mean had been generated. Normalized volumes were then resampled to  $3 \text{ mm}^3$  voxels and convolved with a 9-mm full-width half-maximum (FWHM) isotropic Gaussian kernel. Statistical analysis was conducted in two stages of a mixed-effects model using classical inference [Friston et al., 2002]. Two epoch types corresponding to the experimental manipulation (primed and non-primed conditions) were modeled as effects of interest with delta functions representing each epoch onset, and convolved with a basis function consisting of a single finite impulse response (FIR) encompassing the epoch duration. High- and low-pass filtering were eschewed. Linear contrasts of the parameter estimates of the FIR functions were entered in a group level *t*-test in a second-stage analysis that treated participants as a random effect, and the *t*-values were transformed into corresponding *Z*-scores.

Hypothesis driven regions-of-interest (ROIs) were specified a priori within MNI atlas space using automated anatomical labeling software [Maldjian et al., 2003; Tzourio-Mazoyer et al., 2002] for analyses of data from the OFC (pars orbitalis), inferior prefrontal cortex (pars opercularis), ACC, middle and superior temporal cortices. In the latter three cases, these ROIs were edited manually to

**TABLE II. Descriptive statistics for the two conditions**

Condition	Mean RT	Error (%)
Primed	939.2 (124.9)	3.7
Unprimed	876.1 (85.2)	1.2

Mean response times (RT; standard deviations in parentheses) by participants and percentage errors.



**Figure 3.**

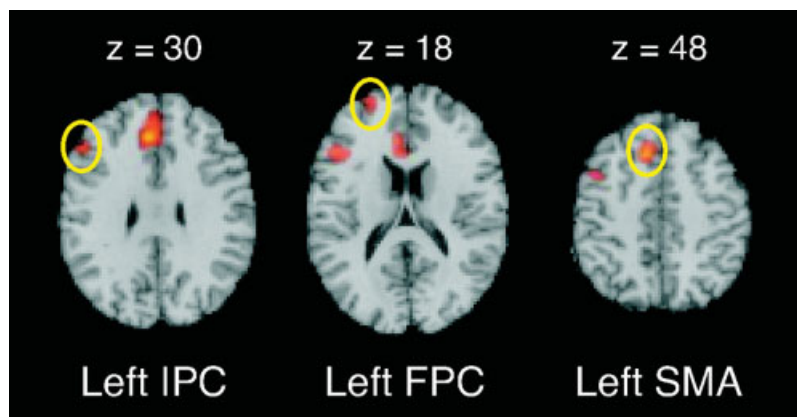
Brain slices showing regions of interest (ROI) and corresponding peak maxima (Max) from the contrast of Primed > Unprimed conditions. PTC, posterior temporal cortex; MTC, mid temporal cortex; ACC, anterior cingulate cortex; pOr, pars orbitalis; pOp, pars opercularis.

conform to the anterior, middle, or posterior y coordinate-based subdivisions specified by Indefrey and Levelt [2004] in their meta-analysis (see their Tables 3 and 4, p. 119). All ROIs were generated for both hemispheres, with the exception of the temporal cortical areas that were generated for the left hemisphere only. Results are reported using  $\alpha$  thresholds of 0.05 small volume corrected (SVC) for multiple comparisons using the false discovery rate (FDR) method [Genovese et al., 2002]. We also conducted whole brain exploratory analyses. These contrasts are reported based upon an  $\alpha$  threshold of 0.001 ( $Z > 3.09$ , uncorrected for multiple comparisons) and a spatial extent threshold of 0.05 (uncorrected; corresponding to 16 or more contiguous voxels in this dataset).

## RESULTS

### Naming Latencies and Error Rates

The analysis of the naming latencies excluded some data. Target-naming responses were discarded if the corresponding prime word was misnamed or omitted. Trials with target-naming errors, omissions, or non-speech sounds were likewise discarded. This resulted in 4.8% of trials overall being excluded from analysis. Mean naming latencies and percentage error rates per condition are shown in Table II. Paired  $t$ -tests were carried out on the participant means from each condition to evaluate the effect of priming a semantic competitor. As expected, primed pictures were named more



**Figure 4.**

Brain slices showing additional cerebral areas activated in the whole brain analysis. Regions circled in yellow were activated in the Primed > Unprimed contrast. IPC, Inferior prefrontal cortex (pars triangularis); FPC, frontopolar cortex; SMA, supplementary motor area.

slowly than pictures that were unprimed,  $t(1,12) = -3.26$ ,  $P < 0.01$ . Percentage error rates (arc-sine transformed) showed no significant effect of condition  $t(1, 12) = 1.76$ ,  $P > 0.05$ .

### Imaging Data

#### Primed versus unprimed contrast

Of the left temporal cortical ROIs examined, only the combined posterior middle and superior temporal gyri showed significant activation for the contrast of primed versus unprimed conditions (Fig. 3). This had a peak in the posterior middle temporal gyrus ( $-60, -45, 6$ ;  $Z = 3.6$ ; Brodmann area [BA] 21). To determine whether there was a trend toward activation in the mid middle temporal ROI, the  $\alpha$  threshold was lowered to 0.001 (uncorrected). This more lenient threshold revealed one significant activation peak ( $-54, -15, -18$ ;  $Z = 3.24$ ; 6 voxels; BA 21). Within the ACC ROIs, significant activation foci were observed in both left ( $-3, 33, 24$ ;  $Z = 5.09$ ; BA 32) and right hemispheres ( $3, 36, 27$ ;  $Z = 4.04$ ; BA 32). This was also the case with the ROIs of the pars orbitalis ( $-45, 33, -15$ ;  $Z = 3.81$ ; BA 11/47, and  $45, 33, -12$ ;  $Z = 3.89$ ; BA 11/47). Significant activation was also found in the left pars opercularis ROI ( $-42, 18, 15$ ;  $Z = 3.36$ ; BA 45). In addition to the regions identified above by the ROI analyses, the whole brain exploratory analysis revealed significant activation foci in the left inferior prefrontal cortex (pars triangularis; BA 46), supplementary motor area (SMA: BA 8/32), and superior frontal gyrus (frontopolar cortex; BA 10; Fig. 4, Table III).

#### Unprimed versus primed contrast

No significant activation was detected in any of the ROIs defined a priori, nor in the whole brain exploratory analysis at the thresholds employed.

## DISCUSSION

In the present experiment, words that were named from a definition and semantically related to a target picture slowed picture-naming latencies compared to unrelated words. This occurred after two intervening trials between the definition and the picture to be named, a demonstration of the competitor-priming effect [Wheeldon and Monsell, 1994]. The fMRI data showed this effect was associated with increased activity in several cortical regions, most of which were predicted a priori, including the left posterior middle temporal gyrus, anterior cingulate, and pars orbitalis of the IFG (considered part of the OFC [Ridderinkhof et al., 2004]).

The results have several implications. First, the activity observed in the left posterior middle temporal gyrus conforms to the region identified by Indefrey and Levelt's [2004] meta-analysis as being involved in word form retrieval. This result is consistent with evidence indicating that electrical stimulation/interruption of this region elicits semantic substitutions during picture naming [e.g., Duffau et al., 2005]. A trend toward increased activity in the midsection of the left middle temporal gyrus was also detected using a more lenient  $\alpha$  threshold ( $P$

$< 0.001$ ).<sup>1</sup> This region conforms to that identified by Indefrey and Levelt [2004] as being involved in lexical-conceptual processing. The finding of activity in both regions is consistent with the results of our earlier fMRI study of picture-word interference effects [de Zubicaray et al., 2001]. We thus interpret the combined temporal cortex activity with the stronger posterior focus as supporting the assumption made by interactive models of spoken word production concerning the locus of competition effects in lexical access [e.g., Harley, 1993]. It is not consistent with the assumption made by serial models that competition occurs and is resolved solely at the lexical-conceptual stage of processing [e.g., Levelt et al., 1999]. If this were the case, then activation increases would have been expected solely in the mid section of the left middle temporal gyrus [Indefrey and Levelt, 2004].

Second, as the competitor priming paradigm intersperses two unrelated trials (one picture and one definition) between the primed distractor word and target picture, an alternative explanation of the posterior temporal activity in terms of concurrent phonological activation occurring in the perceptual network is not tenable [cf. Levelt et al., 1999; Roelofs, 2004]. This explanation was invoked by serial modelers to explain picture-word interference effects involving near simultaneous presentations of target pictures and distractor words [e.g., Cutting and Ferreira, 1999].<sup>2</sup> In order for this explanation to account for the present results, a further assumption and demonstration of refractory phonological activation within the perceptual network would be needed. Explanations of the competitor priming effect have assumed that it is due to the maintenance of high levels of activation among the primed distractor word's features in the production network, increasing competition with the target during subsequent trials [Belke et al., 2005; Wheeldon and Monsell, 1994]. This could be because naming to definition activates a richer (i.e., larger and more diverse) set of features than are activated during standard picture naming [Belke et al., 2005; Vitkovitch et al., 2001]. This account, however, is consistent with the present results only if it is assumed that activation is maintained in terms of both lexical-conceptual units and retrieval of their corresponding word forms [cf. Belke et al., 2005; Wheeldon and Monsell, 1994].

<sup>1</sup> We acknowledge the possibility that this result might represent a false positive. Nebel et al. [2005], however, demonstrated that compressed or sparse imaging sequences such as the one used here achieve slightly less statistical power than do conventional continuous sequences. We therefore consider it reasonable to report this result as indicative of at least a trend.

<sup>2</sup> Levelt et al. [1999] invoked this possibility to permit their serial model a means of explaining Cutting and Ferreira's [1999] results. In that study, participants named pictures of objects with homophone names (e.g., *ball*) whereas auditory distractor words were presented 150 msec before picture onset. Categorically related distractors (e.g., *frisbee*) slowed picture naming, whereas distractors related to the non-depicted meaning (e.g., *dance*) facilitated it. Cutting and Ferreira interpreted this as evidence for interactivity in spoken word production.

**TABLE III. Anatomical labels and coordinates of additional peak activation foci from the whole brain exploratory analysis**

Description	Hemisphere	<i>x</i>	<i>y</i>	<i>z</i>	Z-score
Primed > unprimed conditions					
Inferior frontal gyrus (pars triangularis)	Left	-48	27	30	4.04
Superior frontal gyrus (frontopolar cortex)	Left	-21	57	18	3.50
Supplementary motor cortex	Left	-6	24	48	4.42

Height and cluster extent thresholds  $P < 0.001$  and  $P < 0.05$  (uncorrected), respectively.

Importantly, the extensive activation foci observed in the anterior cingulate and IFG (specifically, the pars orbitalis and pars opercularis) suggest that central processes can and do intervene during lexical access [Ferreira and Pashler, 2002]. There is now considerable evidence from neuroimaging studies indicating that one role of the ACC is monitoring of conflict during information processing [for an overview, see Botwinick et al., 2004]. In fact, the ACC activation that we observed in the right hemisphere was nearly identical to the peak reported by Barch et al [2001] for response conflict effects during vocal task performance (their peak 4, 39, 27; our peak 3, 36, 27). In terms of the competitor priming effect, this can be considered akin to detecting processing bottlenecks during lexical access [Ferreira and Pashler, 2002], possibly due to the activation of mutually inhibiting, competing representations as implemented in some interactive models [e.g., Botwinick et al., 2001; Harley, 1993]. When viewed in conjunction with the finding of increased activity especially in the left posterior middle temporal gyrus, this result may be interpreted as indicating the ACC was detecting competition occurring during (primarily) word form retrieval.

Once conflict is detected by the ACC, control processes are assumed to intervene to guide activation in the language production network to facilitate the current goal (i.e., naming the target picture [de Zubicaray, in press]). Interactive models propose that the distractor's activated features are suppressed by the use of inhibitory links either within or between levels, or by a combination of both [e.g., Berg and Schade, 1992b; Dell and O'Seaghdha, 1994; Harley, 1993]. It therefore seems likely that top-down control is implemented in the form of manipulating the activation levels of competitors through the use of these links. Although some researchers have implicated the pars opercularis of the IFG in inhibitory control, particularly the right hemisphere [e.g., Aron et al., 2004], others have ascribed this role to the OFC [e.g., Fuster, 1997, 2005], which extends laterally into the pars orbitalis [Ridderinkhof et al., 2004]. Both regions showed significant increases in activity in the present study (the former on the left, the latter bilaterally), the latter consistent with evidence that electrical stimulation of this region elicits semantic substitutions during picture naming [e.g., Duffau et al., 2005]. This account necessarily assumes that the speech errors produced during electrical stimulation of the pars orbitalis region represent an interruption of a normal mechanism for competitor deactivation in spoken word production. Rolls [2004] has provided a compelling account of the functions of the OFC in terms of learning and reversal of

stimulus-reinforcement or stimulus-stimulus associations, used in the correction of inappropriate behavioral responses. This account accords well with the notion that control as implemented in the competitor priming paradigm involves weakening or reversing the strengthened association between the primed distractor word and target picture, such that selection of the distractor is suppressed. Notably, Aron et al. [2004] ascribed a similar role of suppressing stimulus-response "mappings" to the pars opercularis (p. 74).

It is worth noting that an alternative account of ACC involvement in speech production was proposed recently by Roelofs [2003] in the context of a serial model. According to this account, the ACC plays a role in setting a task and directing attention to task-relevant goals, rather than conflict detection per se. Although devised to explain Stroop and picture-word interference effects, this account could conceivably be invoked to explain the ACC activation observed in the present study. Eliciting a competing response by naming to definition before naming a target picture thus could be considered to induce a larger need for exerting task-relevant control than would a noncompeting response. As we noted earlier, however, the finding of co-activated lexical-conceptual and word form encoding areas of the left temporal cortex during competitor priming refutes a basic assumption of the serial model approach.

In addition to the regions detected with the ROI analyses, the more lenient whole brain exploratory analysis revealed novel foci showing increased activity during the competitor priming effect. One of these was in the left pars triangularis of the inferior prefrontal cortex, a region that has been proposed to be involved in selecting from among competing conceptual representations [Kan and Thompson-Schill, 2004; Moss et al., 2005]. This finding accords well with the assumption implemented in language production models of competition occurring naturally between lexical-conceptual neighbors, with selection demands increasing with the introduction of a strong lexical-conceptual competitor, as is the case with competitor priming [Dell and Sullivan, 2004; Levelt, 2001]. A number of neuroimaging studies have reported co-activation of the pars triangularis and pars orbitalis during tasks involving semantic processing [for a review, see Bookheimer, 2002]. Another region, frontopolar cortex, is reported regularly in fMRI studies of verbal episodic memory retrieval [e.g., Rugg et al., 2002; de Zubicaray et al., 2005], whereas activity in the left supplementary motor area (SMA) is a common finding in word production studies [Indefrey and Levelt, 2004]. In their meta-analysis, Indefrey and Levelt [2004] reported that activity in the SMA



was detected more reliably during word generation than other production tasks. Accordingly, a conservative interpretation of the above results might be that picture naming after a primed distractor places greater demand on retrieval operations and generative processes than does unprimed picture naming.

Finally, interactivity during spoken word production implies that activated word forms excite lexical-conceptual representations before the act of lexical selection. There are different ways to implement this computationally: Both simple feedback connections and recurrent attractor networks have served this purpose [Dell and Sullivan, 2004; Plaut and McClelland, 1993]. As a number of authors have noted, however, this is a difficult proposition to demonstrate empirically [Dell and Sullivan, 2004; Rumel et al., 2005]. The present finding of increased activity in both lexical-conceptual and word form retrieval-related areas can be interpreted parsimoniously as evidence supporting cascaded processing, in which all (i.e., target and distractor) lexical-conceptual units are activated and their corresponding word forms retrieved, with the latter process finishing last [Dell and Sullivan, 2004]. Interactive models necessarily presuppose cascaded processing [Cutting and Ferreira, 1999; Dell and O'Seaghdha, 2001; Harley, 1993].

In summary, our results provide support for the notion of multiple stages of lexical access during spoken word production, in which activation spreads between stages comprising lexical-conceptual representations and the retrieval of their corresponding word forms [Dell and Sullivan, 2004]. More specifically, they provide support for the assumptions of interactive models concerning the locus of competition effects during lexical access. They also indicate, however, that although competition occurs during lexical access, it is not resolved entirely at this level of processing. Additional central mechanisms of conflict/bottleneck monitoring and control are required. This suggests that lexical selection is biased by top-down mechanisms to select words that meet the current goal of speech.

## ACKNOWLEDGMENTS

We thank Jenny Stephensen, Fiona Mack and Erica Maddock for their assistance. Greig de Zubicaray is supported by an ARC Research Fellowship.

## REFERENCES

- Aguirre G, Zarahn E, D'Esposito M (1998): The variability of human, BOLD hemodynamic responses. *Neuroimage* 8:360–369.
- Aron AR, Robbins TW, Poldrack RA (2004): Inhibition and the right inferior prefrontal cortex. *Trends Cogn Sci* 8:170–177.
- Ashburner J, Friston KJ (1999): Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7:254–266.
- Barch DM, Braver TS, Akbudak E, Conturo T, Ollinger J, Snyder A (2001): Anterior cingulate cortex and response conflict: effects of response modality and processing domain. *Cereb Cortex* 11:837–848.
- Belke E, Meyer AS, Damian MF (2005): Refractory effects in picture naming as assessed in a semantic blocking paradigm. *Q J Exp Psychol A* 58:667–692.
- Berg T, Schade U (1992a): The role of inhibition in a spreading-activation model of language production. I. The psycholinguistic perspective. *J Psycholinguist Res* 21:405–434.
- Berg T, Schade U (1992b): The role of inhibition in a spreading-activation model of language production. II. The simulational perspective. *J Psycholinguist Res* 21:435–462.
- Bonin P, Peereman R, Malardier N, Meacut;ot A, Chalard M (2003): A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behav Res Methods Instrum Comput* 35:158–167.
- Bookheimer S (2002): Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci* 25:151–188.
- Botwinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001): Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652.
- Botwinick MM, Cohen JD, Carter CS (2004): Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* 8:539–546.
- CELEX English database (Release E25) [On-line] (1993): Available: Nijmegen: Centre for Lexical Information [Producer and Distributor]. Available at: <http://www.mpi.nl/world/celex/>. Accessed on: February 18, 2005.
- Cutting JC, Ferreira VS (1999): Semantic and phonological information flow in the production lexicon. *J Exp Psychol Learn Mem Cogn* 25:318–344.
- Damian MF, Martin RC (1999): Semantic and phonological codes interact in single word production. *J Exp Psychol Learn Mem Cogn* 25:345–361.
- de Zubicaray GI (in press): Cognitive neuroimaging: cognitive science out of the armchair. *Brain Cogn*.
- de Zubicaray GI, McMahan KL, Eastburn MM, Finnigan S, Humphreys MS (2005): fMRI evidence of word frequency and strength effects during recognition memory. *Brain Res Cogn Brain Res* 24:587–598.
- de Zubicaray GI, McMahan KL, Eastburn MM, Wilson SJ (2002): Orthographic/phonological facilitation of naming responses in the picture-word task: an event-related fMRI study using overt vocal responding. *Neuroimage* 16:1084–1093.
- de Zubicaray GI, Wilson SJ, McMahan KL, Muthiah S. (2001): The semantic interference effect in the picture-word paradigm: an event-related fMRI study employing overt responses. *Hum Brain Mapp* 14:218–227.
- Dell GS, O'Seaghdha PG (1991): Mediated and convergent lexical priming in language production: a comment on Levelt et al. (1991). *Psychol Rev* 98:604–614.
- Dell GS, O'Seaghdha PG. 1994. Inhibition in interactive activation models of linguistic selection and sequencing. In: Dagenbach D, Carr TH, editors. *Inhibitory processes in attention, memory, and language*. San Diego: Academic Press. p 409–453.
- Dell GS, Sullivan JM. 2004. Speech errors and language production: neuropsychological and connectionist perspectives. In: Ross BH, editor. *The psychology of learning and motivation*. San Diego: Elsevier. p 63–108.
- Duffau H, Gatignol P, Mandonnet E, Peruzzi P, Tzourio-Mazoyer N, Capelle L (2005): New insights into the anatomical-functional connectivity of the semantic system: a study using cortico-subcortical electrostimulations. *Brain* 128:797–810.
- Evans AC, Kamber M, Collins DL, Macdonald D. 1994. An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon S, Fish D, Andermann F, Bydder GM, Stefan H, editors. *NATO ASI series A, Life Sciences: Vol. 264. Magnetic resonance scanning and epilepsy*. New York: Plenum Press. p 263–274.

- Ferreira VS, Griffin ZM (2003): Phonological influences on lexical (mis)selection. *Psychol Sci* 14:86–90.
- Ferreira VS, Pashler H (2002): Central bottleneck influences on the processing stages of word production. *J Exp Psychol [Hum Learn]* 28:1187–1199.
- Forde EME, Humphreys GW (1997): A semantic locus for refractory behaviour: implications for access-storage distinctions and the nature of semantic memory. *Cog Neuropsychol* 14:367–402.
- Freire L, Roche A, Mangin JF (2002): What is the best similarity measure for motion correction in fMRI time series? *IEEE Trans Med Imaging* 21:470–484.
- Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J (2002): Classical and Bayesian inference in neuroimaging: applications. *Neuroimage* 16:484–512.
- Fuster JM. 1997. *The prefrontal cortex—anatomy, physiology, and neuropsychology of the frontal lobe*. Philadelphia: Lippincott-Raven.
- Fuster JM. (2005): The cortical substrate of general intelligence. *Cortex* 41:228–229.
- Genovese CR, Lazar NA, Nichols T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:870–878.
- Gracco VL, Tremblay P, Pike B (2005): Imaging speech production using fMRI. *Neuroimage* 26:294–301.
- Harley TA (1993): Phonological activation of semantic competitors during lexical access in speech production. *Lang Cogn Process* 8:291–309.
- Harley TA, MacAndrew SGB (2001): Constraints upon word substitution speech errors. *J Psycholinguist Res* 30:395–418.
- Kan IP, Thompson-Schill SL (2004): Selection from perceptual and conceptual representations. *Cogn Affect Behav Neurosci* 4:466–482.
- Indefrey P, Levelt WJM. 2000. The neural correlates of language production. In: Gazzaniga M, editor. *The new cognitive neurosciences*. Cambridge, MA: MIT Press. p 845–865.
- Indefrey P, Levelt WJM (2004): The spatial and temporal signatures of word production components. *Cognition* 92:101–144.
- Levelt WJM. (2001). Spoken word production: a theory of lexical access. *Proc Natl Acad Sci USA* 98:13464–13471.
- Levelt WJM, Roelofs A, Meyer AS (1999): A theory of lexical access in speech production. *Behav Brain Sci* 22:1–75.
- Maldjian JA, Laurienti PJ, Burdette JB, Kraft RA (2003): An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19:1233–1239.
- McMahon K, Pringle A, Eastburn M, Maillet D (2004): Improving EPI imaging quality and sound levels with bandwidth selection. *Proceedings of the 12th Annual Meeting of the International Society for Magnetic Resonance in Medicine*, Kyoto, 1033.
- Miozzo M, Caramazza A (2003): When more is less: a counterintuitive effect of distractor frequency in the picture-word interference paradigm. *J Exp Psychol Gen* 132:228–258.
- Morsella E, Miozzo M (2002): Evidence for a cascade model of lexical access in speech production. *J Exp Psychol Learn Mem Cogn* 28:555–563.
- Moss HE, Abdallah S, Fletcher P, Bright P, Pilgrim L, Acres K, Tyler LK (2005): Selecting among competing alternatives: selection and retrieval in the left inferior frontal gyrus. *Cereb Cortex* 15:1723–1735.
- Nebel K, Stude P, Wiese H, Müller B, de Greiff A, Forsteing M, Diener HC, Keidel M (2005): Sparse imaging and continuous event-related fMRI in the visual domain: a systematic comparison. *Hum Brain Mapp* 24:130–143.
- Peterson RR, Savoy P (1998): Lexical selection and phonological encoding during language production: evidence for cascaded processing. *J Exp Psychol Learn Mem Cogn* 24:539–557.
- Plaut DC, McClelland JL. 1993. *Generalization with componential attractors: word and nonword reading in an attractor network*. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates. p 824–829.
- Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS (2004): Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn* 56:129–140.
- Roelofs A (2003): Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychol Rev* 110:88–125.
- Roelofs A (2004): Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: comment on Rapp and Goldrick (2000). *Psychol Rev* 111:561–572.
- Rolls ET (2004): The functions of the orbitofrontal cortex. *Brain Cogn* 55:11–29.
- Rugg MD, Otten LJ, Henson RN (2002): The neural basis of episodic memory: evidence from functional neuroimaging, *Philos Trans R Soc Lond B Biol Sci* 357:1097–1110.
- Ruml W, Caramazza A, Capasso R, Miceli G (2005): Interactivity and continuity in normal and aphasic language production, *Cogn Neuropsychol* 22:131–168.
- Snodgrass JG, Vanderwart M (1980): A standardized set of 260 pictures: norms for naming agreement, familiarity, and visual complexity. *J Exp Psychol [Hum Learn]* 6:174–215.
- Starreveld PA, La Heij W (1996): Time course analysis of semantic and orthographic context effects in picture naming. *J Exp Psychol Learn Mem Cogn* 22:896–918.
- Stroop JR (1935): Studies of interference in serial verbal reactions. *J Exp Psychol* 12:643–662.
- Szekely A, Jacobsen T, D’Amico S, Devescovi A, Andonova E, Herron D, Lu CC, Pechmann T, Pleh C, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Hsu J, Iyer G, Kohnert K, Mehotcheva T, Orozco-Figueroa A, Tzeng A, Tzeng O, Arevalo A, Vargha A, Butler AC, Buffington R, Bates E (2004): A new on-line resource for psycholinguistic studies. *J Mem Lang* 51:247–250.
- Tree JJ, Hirsh KW (2003): Sometimes faster, sometimes slower: associative and competitor priming in picture naming with young elderly participants. *J Neurolinguistics* 16:489–514.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- Vaughan JT, Adriany G, Garwood M, Yacoub E, Duong T, Delabarre L, Andersen P, Ugurbil K (2002): Detunable transverse electromagnetic (TEM) volume coil for high-field NMR. *Magn Reson Med* 47:990–1000.
- Vitkovitch M, Rutter C, Read A (2001): Inhibitory effects during object name retrieval: The effect of interval between prime and target on picture naming responses. *Br J Psychol* 92:483–506.
- Wheeldon L (2003): Inhibitory form priming of spoken word production. *Lang Cogn Process* 18:81–109.
- Wheeldon LR, Monsell S (1994): Inhibition of spoken word production by priming a semantic competitor. *J Mem Lang* 33:332–356.
- Zaitsev M, Hennig J, Speck O (2003): Automated online EPI distortion correction for fMRI applications. *Proceedings of the 11th Annual Meeting of the International Society for Magnetic Resonance in Medicine*, 1042.
- Zeng H, Constable RT (2002): Image distortion correction in EPI: comparison of field mapping with point spread function mapping. *Magn Reson Med* 48:137–146.