The Semantic Interference Effect in the Picture-Word Paradigm: An Event-Related fMRI Study Employing Overt Responses

Greig I. de Zubicaray,* Stephen J. Wilson, Katie L. McMahon,
and Santhi Muthiah

Centre for Magnetic Resonance, The University of Queensland, QLD, Australia

Abstract: We used event-related functional magnetic resonance imaging (fMRI) to investigate neural responses associated with the semantic interference (SI) effect in the picture-word task. Independent stage models of word production assume that the locus of the SI effect is at the conceptual processing level (Levelt et al. [1999]: *Behav Brain Sci* 22:1–75), whereas interactive models postulate that it occurs at phonological retrieval (Starreveld and La Heij [1996]: *J Exp Psychol Learn Mem Cogn* 22:896–918). In both types of model resolution of the SI effect occurs as a result of competitive, spreading activation without the involvement of inhibitory links. These assumptions were tested by randomly presenting participants with trials from semantically-related and lexical control distractor conditions and acquiring image volumes coincident with the estimated peak hemodynamic response for each trial. Overt vocalization of picture names occurred in the absence of scanner noise, allowing reaction time (RT) data to be collected. Analysis of the RT data confirmed the SI effect. Regions showing differential hemodynamic responses during the SI effect included the left mid section of the middle temporal gyrus, left posterior superior temporal gyrus, left anterior cingulate cortex, and bilateral orbitomedial prefrontal cortex. Additional responses were observed in the frontal eye fields, left inferior parietal lobule, and right anterior temporal and occipital cortex. The results are interpreted as indirectly supporting interactive models that allow spreading activation between both conceptual processing and phonological retrieval levels of word production. In addition, the data confirm that selective attention/response suppression has a role in resolving the SI effect similar to the way in which Stroop interference is resolved. We conclude that neuroimaging studies can provide information about the neuroanatomical organization of the lexical system that may prove useful for constraining theoretical models of word production. *Hum. Brain Mapping* 14:218–227, 2001. © 2001 Wiley-Liss, Inc.

Key words: fMRI; semantic interference; word production; middle temporal gyrus; superior temporal gyrus; orbitomedial prefrontal cortex

INTRODUCTION

A considerable amount of research on spoken word production has involved the measurement of object and word naming latencies [Glaser, 1992; Levelt, 1999]. One task, the picture-word paradigm [Rosinski et al., 1975], has been employed extensively in experimental psychology studies to investigate hypotheses regarding the processes involved in word production.

Contract grant sponsor: Viertel Foundation.

*Correspondence to: Greig I. de Zubicaray, Centre for Magnetic Resonance, The University of Queensland, QLD, 4072. Australia.

E-mail: greig.dezubicaray@cmr.uq.edu.au

Received for publication 7 February 2001; accepted 6 August 2001

The paradigm, a generalization of the Stroop [1935] interference task, requires participants to name a target picture while ignoring an accompanying word [for a review of Stroop-like interference effects, see MacLeod, 1991].

A well-documented effect on the picture-word task is that naming latencies are longer when the distractor word is semantically-related to the depicted object than when it is an unrelated word or letter-string [Lupker, 1979; Starreveld and La Heij, 1996]. This semantic interference (SI) effect is restricted to a small range of stimulus onset asynchronies (SOAs) from about -200 msec (word presented first) to +200 msec (picture first) and is at a maximum when target and distractor are presented simultaneously [Glaser and Döngelhoff, 1984; Starreveld and La Heij, 1996]. The component processes of theoretical models of word production and their interactions are continuing sources of debate [Caramazza, 1997; Caramazza and Costa, 2000; Foygel and Dell, 2000; Levelt et al., 1999; Roelofs, 1997; Starreveld and La Heij, 1996, 1999a,b]. The models that do attempt to explicitly account for the effects observed in the picture-word task, however, assume that the locus of the SI effect is either at the level of processing lexical concepts (or more specifically at the "lemma" node containing information about syntactic properties for each lexical concept) [Levelt et al., 1999; Roelofs, 1997] or at the level of retrieving the phonological representation of the target picture's name [Starreveld and La Heij, 1996, 1999a].

Although lexical selection is assumed to occur in word production models, these models do not elaborate upon the allocation of attention for resolving the SI effect [Carr, 1999; Levelt et al., 1999; Starreveld and La Heij, 1996]. Most models of Stroop interference propose an important role for selective attention in its resolution, either through active suppression of distractor information or selective enhancement of target information [MacLeod, 1991]¹. Conversely, most theoretical models of word production do not include inhibitory links within or between levels [Levelt et al., 1999; Starreveld and La Heij, 1996]. Instead, selection is achieved through a process of competitive, spreading activation of both target and related, non-target nodes, with the node with the highest activation being

selected (via the so-called "Luce ratio" in the case of the WEAVER++ model [Levelt et al., 1999]). In the picture-word task, the semantically-related distractor word is considered to enhance the activation of an already coactivated non-target node. The SI effect therefore represents the additional processing time needed to resolve the competition occurring at either the conceptual [Levelt et al., 1999; Roelofs, 1997] or phonological retrieval level [Starreveld and La Heij, 1996]. Some evidence, however, indicates that selective attention may play a role in the resolution of the SI effect though suppression of distractor information, similar to the way in which Stroop interference is resolved [Carr, 1999; Kingma et al., 1996]. Lupker [1982] has also proposed that "the tendency to produce the (distractor) word's name must still be suppressed before the correct response can be emitted."

Relatively few functional imaging studies have attempted to draw inferences about the cognitive architecture mediating word production. In fact, many cognitive scientists remain divided about the utility of functional imaging for answering questions about the modular organization of the cognitive system [for a review of these issues see Bub, 2000]. Information regarding the neuroanatomical organization of the lexical system may prove useful, however, for constraining theoretical models of word production [Caramazza, 1997; Humphreys et al., 1999; Shelton and Caramazza, 1999]. In their extensive review of 58 neuroimaging experiments involving picture naming and word generation tasks, Indefrey and Levelt [2000] concluded that the core process of word production is subserved by a left-lateralized network involving the posterior inferior frontal gyrus (Broca's area), the middle section of the superior and middle temporal gyri, the posterior superior and middle temporal gyri (Wernicke's area) and the left thalamus. The conclusions of particular relevance to the present study were that the mid section of the middle temporal gyrus appeared to be involved in conceptual processing, and that Wernicke's area was specifically involved in phonological code retrieval [Indefrey and Levelt, 2000].

The time-course of the SI effect in the picture-word task is well-documented [Glaser and Döngelhoff, 1984; Starreveld and La Heij, 1996]. The underlying neural mechanisms, however, are not. Therefore, in the present study, we examined the neural substrates of the SI effect in the picture-word task using event-related functional magnetic resonance imaging (fMRI). An important feature of the event-related method is that it allows trials to be randomized, whereas blocked designs may result in participants employing anticipation strategies that allow them to

¹Although the SI effect is considered a form of Stroop-like interference [MacLeod, 1991], computational models of Stroop interference cannot directly simulate the SI effect in the picture-word task as they do not include representational structures for semantically related and unrelated words [Starreveld and La Heij, 1996].

be optimally prepared for the trials in a given condition [D'Esposito et al., 1999], an issue that has been raised previously in relation to behavioral studies employing the picture-word paradigm [Lupker, 1982; Starreveld and La Heij, 1996]. Subvocal articulation, although employed to avoid movement-related artifacts in the majority of fMRI studies of word production [Hickok et al., 2000], cannot be used as a simple substitute for overt vocalization due to differences in evoked hemodynamic response patterns between overt and subvocal conditions [Barch et al., 1999; Rosen et al., 2000]. In addition, scanner noise resulting from the rapid switching of gradients can be problematic for fMRI studies of language processing as it can activate the superior and middle temporal gyri as well as primary auditory cortex [MacSweeney et al., 2000]. The event-related design employed in the present study avoided these potential confounds by interleaving behavioral trials with the sparse acquisition of images, thus permitting overt vocalization and excluding the effects of scanner noise [Eden et al., 1999; Elliott et al., 1999].

This combination of sparse event-related fMRI with overt responding should allow predictions regarding the cognitive architecture mediating the SI effect to be tested. Specifically, if the locus of the SI effect in the picture-word task is at the conceptual processing level [Levelt et al., 1999] then differential hemodynamic responses associated with this effect can be predicted in the left mid-section of the middle temporal gyrus based upon the results of Indefrey and Levelt's [2000] review. Alternatively, if the locus of the SI effect is at the level of retrieving phonological codes [Starreveld and La Heij, 1996], then differential responses can be predicted in the left posterior superior and middle temporal gyri. In addition, if selective attention is involved in the resolution of the SI effect, differential responses can be predicted in a network of cerebral areas similar to those identified as being associated with the resolution of Stroop interference in functional imaging studies that have employed *overt* responding, including the anterior cingulate cortex, the orbital and medial prefrontal cortex, and posterior parietal cortex [Barch et al., 1999; Bench et al., 1993; Carter et al., 1995; Peterson et al., 1999].

MATERIALS AND METHODS

Participants

Eight right-handed participants (four female) took part in the study. Mean age was 28.6 years ($SD = 1.8$). All were native English speakers. No participant had a

history of psychiatric or neurological disorder, nor history of head trauma or substance abuse. Informed consent was gained from all participants before the commencement of the study. The study was conducted within the constraints of the ethical clearance from the Medical Research Ethics Committee of the University of Queensland for MRI experiments on humans at the Centre for Magnetic Resonance. Participants received a small gratuity and an MR image of their brain for their involvement in the study.

Picture-Word Paradigm

Forty line-drawings from the Snodgrass and Vanderwart [1980] corpus of pictures were used as target stimuli. The line drawings were taken from a number of semantic categories: animals, vegetables, body parts, furniture, musical instruments, tools, clothing, vehicles, and insects. Each picture served in two distractor conditions: 1) *semantic interference*: distractor words were taken from the same semantic category as the target picture; and 2) *control*: a row of Xs comprising the number of letters in each depicted object's name. For example, the following distractors accompanied the picture of a *mouse*: *dog* (semantically-related) and XXXXX (control). Target picture names and semantically-related distractor words were matched for word length (target $M = 5.28$; semantically-related $M = 5.25$, $F < 1$) and number of syllables (target $M = 1.58$; semantically-related $M = 1.70$, $F < 1$).

Immediately before the imaging experiment each participant underwent a practice session outside of the magnet to train them to use the correct names for the 40 target stimuli. This involved three consecutive presentations of the target pictures on a laptop computer. In the first two presentations, the participants saw all the target pictures together with the correct names printed just below the pictures. Participants were instructed to study the pictures and name them, using the printed words if necessary. Any erroneous responses were corrected by the experimenter. In the final presentation, they were instructed to name the pictures as quickly as possible without the correct names printed below them.

For the imaging session, each trial consisted of a target picture and distractor displayed for a period of 750 msec ($SOA = 0$ sec) followed by a blank screen for 3.25 sec while the participant made a vocal response, and an 11 sec interstimulus interval (ISI) during which participants viewed a fixation point (+) and an image volume was acquired (see Fig. 1 and image acquisition details below). Participants were instructed to name

Figure 1.

(A) A single trial from the event-related experimental design and (B) recording of vocal responses.

the pictures as quickly as possible while ignoring the printed words, and not to speak or move during imaging (as indicated by the scanner noise). The trials from the two distractor conditions were presented in random order and randomized between participants to prevent order effects. A total of 80 trials were presented over a 20 min imaging session (40 pictures $\times$ two distractor conditions). The size of the pictures including background was approximately 10×10 cm. Distractors were presented in lower-case 36 point Times New Roman font and located in such a manner as to allow the identification of both picture and distractor with minimal difficulty. This location was held constant throughout the experiment. The line drawings and distractors were presented in black on a luminous white background, enlarged and back-projected onto the centre of a screen at the foot of the bore of the magnet. Participants viewed the projection screen through a mirror mounted at the top of the head coil. The stimuli subtended approximately 10° of visual arc when the participant was positioned for scanning. Vocal responses and reaction times were recorded on digital audio files via an MR compatible microphone attached to the head coil.

Image Acquisition

Images were acquired using a 2 Tesla Bruker Medspec S200 system at the Centre for Magnetic Resonance, Brisbane. A quadrature Helmholtz head coil was used for RF reception. For the functional MRI studies, 80 T_2^* -weighted gradient echo echoplanar images (EPI) depicting blood oxygen level dependent (BOLD) contrast [Ogawa et al., 1990] were acquired

with an interleaved sequence in each of 21 planes parallel to the anterior-posterior commissure with TE 38 msec, effective TR 15,000 msec (see below), in-plane resolution 3.44 mm and slice thickness 5 mm (zero gap). Behavioral trials were interleaved with detection of BOLD signal changes using a sparse acquisition method [Eden et al., 1999; Elliott et al., 1999] that utilized the delay between task-related neural activity and its estimated peak hemodynamic response (time-to-peak approximately 4.7 ± 1.1 sec) [Aguirre et al., 1998]. For each trial, no field gradients were applied for a 4 sec period of relative silence allowing for stimulus presentation and the participant's overt verbal response, then immediately applied for image acquisition (see Fig. 1). A single image volume was acquired within 3 sec, approximately coincident with the trial's estimated peak hemodynamic response. No field gradients were applied for an additional 8 sec period to allow the BOLD response to the scanner noise to return to baseline [Elliott et al., 1999]. Total imaging time was 20 min for the 80 trials. Head movement was limited by foam padding within the head coil and a restraining band across the forehead. In the same session, a high-resolution 3D T_1 image was acquired using an MP-RAGE sequence with TI 850 msec, TR 1300 msec, TE 5.2 msec, and slice resolution 0.9 mm^3 .

Image Analysis

Image processing and statistical analyses were carried out using statistical parametric mapping software (SPM99; Wellcome Department of Cognitive Neurology, Queens Square, London, UK). Each participant's T_2^* -weighted images were realigned to the first image of the series using a rigid-body transformation procedure and a mean image created from the realigned data [Friston et al., 1995]. The high-resolution T_1 and mean T_2^* -weighted images were then spatially normalized via non-linear basis functions to the T_1 and EPI template images, respectively, included in SPM99 [Ashburner and Friston, 1999]. These templates conform to the space defined by the International Consortium for Brain Mapping project (ICBM) (NIH P-20 grant) and closely approximate the stereotaxic space described in the atlas of Talairach and Tournoux [1988] [Evans et al., 1994]. The normalized T_1 images were then segmented using the Brain Extraction Tool (BET) [Smith, 2000]. The non-linear transformations for the mean T_2^* -weighted images were subsequently applied to the realigned time-series data. After spatial normalization, the fMRI datasets were identically smoothed with an isotropic Gaussian kernel (full

width half maximum [FWHM] = 10 mm) for the group analysis to increase signal to noise and to accommodate variability in gyral anatomy and error of voxel displacement during normalization. A general linear model was applied to the time-course of activation of each voxel [Friston et al., 1995; Worsley and Friston, 1995]. The model included separate basis functions (non-windowed first-order Fourier series) for trial-specific BOLD responses to correct trials from the two distractor conditions. Non-windowed Fourier series have the ability to fit any responses within the frequency range specified, and have the advantage of being phase invariant. The temporal high-pass filter was set to the image acquisition time (TA) of 3 sec, as confounds attributable to low frequency drift in the BOLD signal were considered unlikely to be present beyond this period. Low-pass filtering was not employed. Basis functions consisting of a synthetic hemodynamic response function were not chosen as each T_2^* -weighted image was acquired during the estimated peak hemodynamic response associated with each trial [Josephs et al., 1997; Josephs and Henson, 1999]. We compared lexical control and semantically-related distractor conditions in a fixed effects analysis with a contrast that tested the sum of pairwise differences between Fourier components [Friston et al., 1999]. This resulted in an F -statistic for every voxel. Statistical inferences were based on the theory of random Gaussian fields. In the case of the superior and middle temporal gyri, anterior cingulate and orbitomedial prefrontal cortex in which we had region-specific hypotheses, correction for multiple comparisons was based on the volume of interest and the smoothness of the underlying SPM [Worsley et al., 1996]. For other brain regions, the correction was for multiple comparisons over the entire brain volume. The threshold adopted for all analyses was $P < 0.05$ (corrected).

RESULTS

Picture Naming Data

Each participant's digital audio file was scored for response accuracy and reaction times measured using voice-key software. Incorrect responses and trials in which participants failed to provide a response were removed from the analysis. Reaction times larger than 2,500 msec were also removed. Table I shows the participants' mean reaction times in milliseconds and error percentages per condition. An analysis of variance (ANOVA) was conducted on the reaction time data that showed a significant effect of distractor condition, $F(1,622) = 8.04$, $P < 0.01$. Error rates for the two

TABLE I. Participants' mean reaction times in milliseconds and error percentages per condition

Distractor condition	Mean RT	% error
Semantically-related	1128	3.4
Lexical control	1067	1.3

distractor conditions were considered insufficient to conduct a meaningful analysis.

Image Data

As per the behavioral data, images for trials involving incorrect or omitted responses were removed before analysis. Relative to the lexical control condition, the semantically-related distractor condition produced differential hemodynamic responses in the left mid- and right anterior middle temporal gyrus (Brodmann area [BA] 21), left posterior superior temporal gyrus (BA 13), left anterior cingulate cortex (BA 32), and orbitomedial prefrontal cortex (OMPFC) comprising right medial and left superior frontal gyri (BA 10). Further differential responses were observed in the right superior frontal gyrus in the region of the frontal eye fields (BA 8), left inferior parietal lobule (BA 40) and right occipital cortex (BA 17). The coordinates and statistical significance of these foci are summarized in Table II.

Figure 2 shows the location of the left superior and middle temporal gyrus differential responses, superimposed on sagittal slices from the group mean high-resolution image. The responses in the left anterior cingulate are shown on a sagittal slice in Figure 3 along with the locations of the OMPFC differential responses displayed on a coronal slice.

DISCUSSION

The present study represents the first event-related fMRI study using overt vocalization to identify differential hemodynamic responses associated with performance on the picture-word task. Our results provide evidence that several brain regions exhibit differential responses associated with the SI effect. This combination of event-related fMRI with overt responding to provide information about the neuroanatomical organization of the lexical system may prove useful for constraining theoretical models of word production [Foygel and Dell, 2000; Levelt et al., 1999; Starreveld and La Heij, 1996].

Based upon Indefrey and Levelt's [2000] conclusions from their meta-analysis of neuroimaging stud-

TABLE II. Differential activations: comparison of semantically-related and lexical control distractor conditions (SI effect)

Side	Cerebral region	F	P-value	Coordinates		
				x	y	z
L	Middle temporal gyrus	9.39	<.05 ^b	-56	-20	-14
R		13.52	<.05 ^a	62	4	-16
L	Superior temporal gyrus	8.13	<.05 ^b	-46	-46	16
R	Medial frontal gyrus	7.56	<.05 ^b	10	62	2
L	Anterior cingulate	8.96	<.05 ^b	-4	42	-4
L	Superior frontal gyrus	6.71	<.05 ^b	-30	60	0
R		14.94	<.05 ^a	18	28	58
L	Inferior parietal lobule	11.62	<.05 ^a	-54	-34	40
R	Occipital lobe	12.51	<.05 ^a	8	-76	-14

^a P-value for entire brain volume.^b P-value corrected for volume of interest.

ies of word production, we hypothesized that if the locus of the SI effect in the picture-word task is at the level of retrieving phonological codes as some theoretical models assume [Starreveld and La Heij, 1996], then differential hemodynamic responses associated with this effect should be observable in the left posterior superior and middle temporal gyri (MTG). Alternatively, if it is at the conceptual processing level [Levelt et al., 1999; Roelofs, 1997] then differential hemodynamic responses should occur in the left mid-section of the MTG. Instead, we observed differential responses in the left mid-section of the MTG *and* in the left posterior superior temporal gyrus (STG), raising the possibility that the SI effect occurs at *both* conceptual and phonological retrieval levels. Although this

result was unanticipated, it does provide indirect support for so-called “interactive” models of word production that propose bi-directional connections between conceptual and phonological levels of processing (i.e., activation spreads in both directions) [Foygel and Dell, 2000; Starreveld and La Heij, 1996]. It is therefore consistent with a recent chronometric study that supported an interactive account [Damian and Martin, 1999]. The finding of differential responses in both regions, however, is difficult to explain in terms of independent stage or strict serial models of word production that assume the locus of the SI effect is solely at the lexical concept/lemma level [Levelt et al., 1999; Roelofs, 1997]. In these latter models, the phonological retrieval level does not feed

Figure 2.

Differential responses in the left mid section of the middle temporal gyrus (mid MTG) and posterior superior temporal gyrus (STG), superimposed on sagittal slices from the mean high-resolution image of the eight participants in the study.

Figure 3.

Differential responses in the left anterior cingulate (ACC; sagittal slice) and the orbitomedial prefrontal cortex (OMPFC; coronal slice) superimposed on the mean high-resolution image of the eight participants in the study.

back to the lemma level [Levelt, 1999]. The slightly higher F -value for the differential response observed in the mid-MTG, however, may indicate that most of the competition associated with the SI effect does occur at the conceptual processing level as postulated by these models.

In both interactive and serial models of word production, resolution of the SI effect is considered to occur via a process of competitive, spreading activation of target and related, non-target nodes, with the node with the highest activation favored for selection [Levelt et al., 1999; Roelofs, 1997; Starreveld and La Heij, 1996]. No inhibitory links within or between levels are included in these models, nor are any general theories about the allocation of attention [Carr, 1999; Wheeldon, 1999]. As argued in the introduction, however, there is evidence that the resolution of the SI effect may rely at least in part on an attentional selection mechanism similar to that considered to operate in models of Stroop interference [Kingma et al., 1996; MacLeod, 1991]. We hypothesized that if a similar selective attention mechanism was involved in resolving the SI effect, then differential responses would be observed in regions previously identified as being associated with the resolution of Stroop interference in functional imaging studies that employed overt responding [Barch et al., 1999; Bench et al., 1993; Carter et al., 1995; Peterson et al., 1999]. Consistent with this hypothesis, we observed differential responses in the anterior cingulate cortex (ACC), OMPFC and inferior parietal cortex.

In the Peterson et al. [1999] factor analysis of Stroop fMRI data, the left inferior parietal cortex loaded on a

factor considered to represent an attentional subsystem involved in monitoring and updating of task demands and performance. This may also explain the differential responses in the inferior parietal cortex in the present study, as task demands are clearly higher in the semantically related distractor condition. Recent neuroimaging studies have confirmed a role for the ACC in detecting the degree of response conflict or competition elicited by a task and indicating the need to engage other brain regions involved in control processes [Barch et al., 2000; Carter et al., 2000]. According to Fuster et al. [2000], "a most characteristic function of orbitomedial prefrontal cortex is *the inhibitory control of interference*" (p. 240; italics in original). We have similarly observed differential responses in the OMPFC during performance of a task requiring suppression of a semantic category response [de Zubicaray et al., 2000]. In the context of the picture-word task, this suggests that when the interference produced by the distractor word is detected by the ACC, processes for implementing inhibitory control in the OMPFC are engaged and the tendency to read the distractor word is overcome. This accords with Lupker's [1982] earlier proposal that to successfully perform the picture-word task the tendency to produce the distractor word's name needs to be suppressed. Consequently, our data suggest a need for inhibitory links to be included in models of word production [Harley, 1993; Wheeldon, 1999].

We also observed differential responses in the right anterior temporal cortex, occipital cortex and frontal eye field. It seems reasonable to assume that the responses in the latter region reflect a greater demand

for oculomotor scanning of the lexical content of the distractor in the semantically-related condition. Signal increases in the right anterior temporal lobe have been reported previously during word-picture matching in neuroimaging experiments [Moore and Price, 1999], and very similar processes may be expected to be engaged by the picture-word task. The right occipital cortex has a well-established role in object processing [Humphreys et al., 1999]. A related possible explanation for the differential responses in the right temporal and occipital areas is that they represent accessing of visual knowledge about both targets and distractors, or more specifically, competition introduced by semantically as well as perceptually similar distractors (e.g., picture, EAGLE; distractor, OWL) [Humphreys et al., 1999; Moore and Price, 1999]. This might explain the increased SI effect observed in Kingma et al. [1996] of patients with right hemisphere lesions compared to patients with left hemisphere lesions. Thus, there may be two parallel competitive mechanisms contributing to the longer response latencies associated with the SI effect in the picture-word task. Though tentative, this interpretation appears worthy of further research with complementary methods such as magnetoencephalography (MEG) or electroencephalography (EEG) that have superior temporal resolution for resolving the time-courses of different competitive mechanisms.

Some limitations of the experimental design and statistical modelling of responses should also be considered when interpreting our data. For example, we used neutral letter-strings (Xs) rather than unrelated words that are often employed as control stimuli in picture-word experiments (although the former are also frequently utilized). These stimuli typically result in shorter response latencies than unrelated words. This is attributed to the incorrect information that is contained in the latter [Starreveld and La Heij, 1996]. We utilized letter-strings to obtain a maximal SI effect within the limited number of trials permitted by the sparse acquisition method. In addition, as image volumes were acquired after a delay to coincide with the estimated peak hemodynamic response to trials from both the control and semantically-related distractor conditions, one possibility is that we may have failed to detect differential responses at certain levels of word production. Although we cannot exclude this possibility, it seems unlikely as regions considered to be involved in several levels (e.g., object processing, conceptual processing, phonological retrieval, lexical selection) were identifiable in our results [Indefrey and Levelt, 2000]. Our ability to detect these differential responses may be attributable to the relatively increased sensitivity to small signal changes afforded

by sparse acquisitions that use long TR intervals (15 sec in the present study) [Eden et al., 1999]. Finally, although non-windowed Fourier series are relatively unconstrained with respect to detecting responses in a specified time-frame, they do not provide unequivocal information regarding the direction of responses. One reason for adopting this approach was that it allowed the time-courses of the responses to the two conditions to be plotted and compared, unlike simpler analyses (e.g., *t*-tests) that involve averaging responses across the two conditions. Examination of the responses of the foci listed in Table II appeared to indicate that all were signal increases relative to the lexical control condition.

In summary, our results indicate that the SI effect in the picture-word task may occur at both conceptual processing *and* phonological retrieval levels of word production. As such, they may be interpreted as indirectly supporting interactive models that assume bi-directional connections between these levels [Damian and Martin, 1999; Starreveld and La Heij, 1996], rather than strict serial models that do not assume spreading activation between these levels [Roelofs, 1997]. In addition, our data suggest that models of word production need to be modified to explain how the competition occurring at these levels is resolved through allocation of attention. An obvious solution would be to include inhibitory links [Harley, 1993]. Finally, they raise the possibility that an additional competitive mechanism centered on object knowledge may contribute to the longer response latencies observed in the distractor condition. More broadly, the data appear to confirm that neuroimaging studies can provide information about the neuroanatomical organization of the lexical system that may prove useful for constraining theoretical models of word production [Shelton and Caramazza, 1999].

ACKNOWLEDGMENTS

This work was supported by a capital equipment grant from the Viertel Foundation. We would like to thank Graham Galloway for his assistance with the pulse sequence, and Steve Williams and Fernando Zelaya for their comments on an earlier draft of this article.

REFERENCES

- Aguirre GK, Zarahn E, D'Esposito M (1998): The variability of human, BOLD hemodynamic responses. *Neuroimage* 8:360–369.

- Ashburner J, Friston KJ (1999): Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* 7:254–266.
- Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD (1999): Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *Neuroimage* 10:642–657.
- Barch DM, Braver TS, Sabb FW, Noll DC (2000): Anterior cingulate and the monitoring of response conflict: evidence from an fMRI study of overt verb generation. *J Cogn Neurosci* 12:298–309.
- Bench CJ, Frith CD, Grasby PM, Friston KJ, Paulesu E, Frackowiak RSJ, Dolan RJ (1993): Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia* 31:907–922.
- Bub DN (2000): Methodological issues confronting PET and fMRI studies of cognitive function. *Cogn Neuropsychol* 17:467–484.
- Caramazza A (1997): How many levels of processing are there in lexical access? *Cogn Neuropsychol* 14:177–208.
- Caramazza A, Costa A (2000): The semantic interference effect in the picture-word interference paradigm: does the response set matter? *Cognition* 75:B51–B64.
- Carr TH (1999): How does WEAVER pay attention? *Behav Brain Sci* 22:39–40.
- Carter CS, MacDonald A, Botvinick M, Ross LL, Stenger VA, Noll D, Cohen JD (2000): Parsing executive processes: strategic vs. evaluative functions of the anterior cingulate cortex. *Proc Natl Acad Sci USA* 97:1944–1948.
- Carter CS, Mintun M, Cohen JD (1995): Interference and facilitation effects during selective attention: an H₂¹⁵O PET study. *Neuroimage* 2:264–272.
- Damian MF, Martin RC (1999): Semantic and phonological codes interact in single word production. *J Exp Psychol Learn Mem Cogn* 25:345–361.
- D'Esposito M, Zarahn E, Aguirre GK (1999): Event-related functional MRI: implications for cognitive psychology. *Psychol Bull* 125:155–164.
- de Zubicaray GI, Zelaya FO, Andrew C, Williams SCR, Bullmore ET (2000): Cerebral regions associated with verbal response initiation, suppression and strategy use. *Neuropsychologia* 38:1292–1304.
- Eden GF, Joseph JE, Brown HE, Brown CP, Zeffiro TA (1999): Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: the behavior interleaved gradients technique. *Magn Reson Med* 41:13–20.
- Elliott MR, Bowtell RW, Morris PG (1999): The effect of scanner sound in visual, motor, and auditory functional MRI. *Magn Reson Med* 41:1230–1235.
- Evans AC, Kamber M, Collins DL, Macdonald D (1994): An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon S, Fish D, Andermann F, Wydder GM, Stefan H, editors. *NATO ASI series A, Life Sciences: Vol. 264. Magnetic resonance scanning and epilepsy*. New York: Plenum. p. 263–274.
- Foygel D, Dell GS (2000): Models of impaired lexical access in speech production. *J Mem Lang* 43:182–216.
- Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ (1995): Spatial registration and normalization of images. *Hum Brain Mapp* 2:165–189.
- Friston KJ, Holmes AP, Poline J-B, Grasby PJ, Williams SC, Frackowiak RSJ, Turner R (1995): Analysis of fMRI time-series revisited. *Neuroimage* 2:45–53.
- Friston KJ, Holmes AP, Worsely KJ (1999): How many subjects constitute a study? *Neuroimage* 10:1–5.
- Fuster JM, Van Hoesen GW, Morecraft RJ, Semendeferi K (2000): Executive systems. In: Fogel BS, Schiffer RB, Rao SM, editors. *Synopsis of neuropsychiatry*. Philadelphia, PA: Lippincott, Williams and Wilkins. p. 229–242.
- Glaser WR (1992): Picture naming. *Cognition* 42:61–105.
- Glaser WR, Döngelhoff F-J (1984): The time-course of picture-word interference. *J Exp Psychol Hum Percept Perform* 10:640–654.
- Harley TA (1993): Phonological activation of semantic competitors during lexical access in speech production. *Lang Cogn Proc* 8:291–309.
- Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, Strick PL, Ugurbil K (2000): A functional magnetic resonance imaging study of the role of the left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. *Neurosci Lett* 287:156–160.
- Humphreys GW, Price CJ, Riddoch MJ (1999): From objects to names: a cognitive neuroscience approach. *Psychol Res* 62:118–130.
- 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.
- Josephs O, Henson RN (1999): Event-related functional magnetic resonance imaging: modelling, inference and optimization. *Philos Trans R Soc Lond B Biol Sci* 354:1215–1228.
- Josephs O, Turner R, Friston KJ (1997): Event-related fMRI. *Hum Brain Mapp* 5:243–248.
- Kingma A, La Heij W, Fasotti L, Eling P (1996): Stroop interference and disorders of selective attention. *Neuropsychologia* 34:273–281.
- Levelt WJM (1999): Models of word production. *Trends Cogn Sci* 3:223–232.
- Levelt WJM, Roelofs A, Meyer AS (1999): A theory of lexical access in speech production. *Behav Brain Sci* 22:1–75.
- Lupker SJ (1979): The semantic nature of response competition in the picture-word interference task. *Mem Cogn* 7:485–495.
- Lupker SJ (1982): The role of phonetic and orthographic similarity in picture-word interference. *Can J Psychol* 36:349–367.
- MacLeod CM (1991): Half a century of research on the Stroop effect: an integrative review. *Psychol Bull* 109:163–203.
- MacSweeney M, Amaro E, Calvert GA, Campbell R, David AS, McGuire P, Williams SCR, Woll B, Brammer MJ (2000): Silent speech reading in the absence of scanner noise: an event-related fMRI study. *Neuroreport* 11:1729–1733.
- Moore CJ, Price CJ (1999): A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122:943–962.
- Ogawa S, Lee TM, Kay AR, Tank DW (1990): Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87:9868–9872.
- Peterson BS, Skudlarski P, Gatenby JC, Zhang H, Anderson AW, Gore JC (1999): An fMRI study of Stroop word-color interference: evidence for cingulate subregions subserving multiple distributed attentional systems. *Biol Psych* 45:1237–1258.
- Roelofs A (1997): The WEAVER model of word-form encoding in speech production. *Cognition* 64:249–284.
- Rosen HJ, Ojemann JG, Ollinger JM, Petersen SE (2000): Comparison of brain activation during word retrieval done silently and aloud using fMRI. *Brain Cogn* 42:201–217.
- Rosinski RR, Golinkoff RM, Kukish KS (1975): Automatic semantic processing in a picture-word interference task. *Child Dev* 46:247–253.

- Shelton JR, Caramazza A (1999): Deficits in lexical and semantic processing: implications for models of normal language. *Psychon Bull Rev* 6:5–27.
- Smith SM (2000): Robust automated brain extraction. *Neuroimage* 11:S625.
- Snodgrass JG, Vanderwart M (1980): A standardized set of 260 pictures: norms for naming agreement, familiarity, and visual complexity. *J Exp Psychol Hum Learn Mem* 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.
- Starreveld PA, La Heij W (1999a): Word substitution errors in a speeded picture-word task. *Am J Psychol* 112:521–553.
- Starreveld PA, La Heij W (1999b): What about phonological facilitation, response-set membership, and phonological coactivation? *Behav Brain Sci* 22:56–58.
- Stroop JR (1935): Studies of interference in serial verbal reactions. *J Exp Psychol* 18:643–662.
- Wheeldon LR (1999): Competitive processes during word-form encoding. *Behav Brain Sci* 22:59–60.
- Worsley KJ, Friston KJ (1995): Analysis of fMRI time-series revisited—again. *Neuroimage* 2:173–181.
- Worsley KJ, Marret S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996): A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73.