

## JUDGING SEMANTIC SIMILARITY: AN EVENT-RELATED fMRI STUDY WITH AUDITORY WORD STIMULI

B. Z. MAHON<sup>a,b,\*</sup> AND A. CARAMAZZA<sup>a,c</sup>

<sup>a</sup>Center for Mind/Brain Sciences (CIMEC), University of Trento, Trento, Italy

<sup>b</sup>Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA

<sup>c</sup>Department of Psychology, Harvard University, Cambridge, MA, USA

**Abstract**—Much of mental life consists in thinking about object concepts that are not currently within the scope of perception. The general system that enables multiple representations to be maintained and compared is referred to as “working memory” [Repovš G, Baddeley A (2006) *Neuroscience* 139:5–21], and involves regions in medial and lateral parietal and frontal cortex [e.g., Smith EE, Jonides J (1999) *Science* 283:1657–1661]. It has been assumed that the contents of working memory index information in regions of the brain that are critical for processing and storing object knowledge. To study the processes involved in thinking about common object concepts, we used event related fMRI to study BOLD activity while participants made judgments of conceptual similarity over pairs of sequentially presented auditory words. Through a combination of conventional fMRI analysis approaches and multi-voxel pattern analysis (MVPA), we show that the brain responses associated with the second word in a pair carry information about the conceptual similarity between the two members of the pair. This was the case in frontal and parietal regions involved in the working memory and decision components of the task for both analysis approaches. However, in other regions of the brain, including early visual regions, MVPA permitted classification of semantic distance relationships where conventional averaging approaches failed to show a difference. These findings suggest that diffuse and statistically sub-threshold “scattering” of BOLD activity in some regions may carry substantial information about the contents of mental representations. © 2010 IBRO. Published by Elsevier Ltd. All rights reserved.

**Key words:** fMRI, conceptual distance, semantic memory, working memory, multi-voxel pattern analysis.

The study of the organization and representation of object knowledge in the human brain has a long tradition in cognitive science, and more recently, has been the focus of numerous functional imaging studies (for review, see Martin, 2007). Many functional imaging studies of semantic memory have focused on understanding how the brain responds to different object concepts (e.g., tools, faces, houses) when they are presented to participants in the form of pictures. However,

much of our mental life involves thinking about object concepts that are not, when we think about them, also the objects of perception. Spoken language is one special case of this situation, where auditory information is mapped onto conceptual knowledge about the referent of a word. Thus, auditorily presented words may be used as stimuli for probing the types of knowledge that are retrieved when thinking about objects, in the absence of any visual input.

In the current experiment, we were interested in studying the role played by the similarity between two different concepts being held in working memory in shaping the brain's response to those concepts (see Cantlon et al., 2009, for review and discussion). A set of materials was selected such that every concept (e.g., “Chair”) was relatively close (i.e., conceptually similar) to another item in the set (“Stool”) and relatively far (i.e., dissimilar) from another concept (e.g., “Stove”). Thus, materials were selected in pairs, such that there was always another concept within the set that was relatively close; the conceptually “far” condition was created by repairing each word with another item from the same broad semantic class. A third level of conceptual distance was the “Identity” condition (every word paired with itself). Conceptual distance (as manipulated between the close and far conditions) was defined at the time of stimulus selection through a combination of intuition and pilot work. However, our measure of conceptual distance is operationalized in terms of participants' actual judgments about the “similarity of the concepts,” obtained during fMRI scanning (see below).

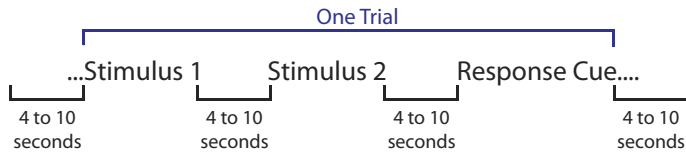
Event-related fMRI was used while participants made judgments of conceptual similarity over the pre-selected pairs of auditorily presented words. As described above, the two words presented on each trial could have three levels of conceptual similarity: they could be the same (“Identity” Condition, e.g., Chair–Chair), conceptually similar (“Close” Condition, e.g., Stool–Chair) or relatively dissimilar (“Far” Condition, e.g., Stove–Chair) (see Fig. 1A for a schematic of the trial structure). Across the whole scanning session, and for every participant, the same pairs of words appeared as the first and second members of the pairs (i.e., stimulus 1 (S1) and stimulus 2 (S2), respectively), and at all relative distances to the other members of the pairs. In this way, the psycholinguistic properties of the words were “matched” across all conditions of interest (S1, and then within S2, Identity, Close, Far) because the same words appeared in all conditions, just repaired in such a way as to derive the desired manipulation of semantic distance. This design permits a direct test of whether the brain responses associated with processing the second object concept in a pair are modulated as a function of the immediately preceding context (i.e., the first object concept of the pair). The critical manipulation in this regard is

\*Correspondence to: B. Z. Mahon, Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627-0268, USA. Tel: +1-585-276-5234; fax: +1-585-276-2127.

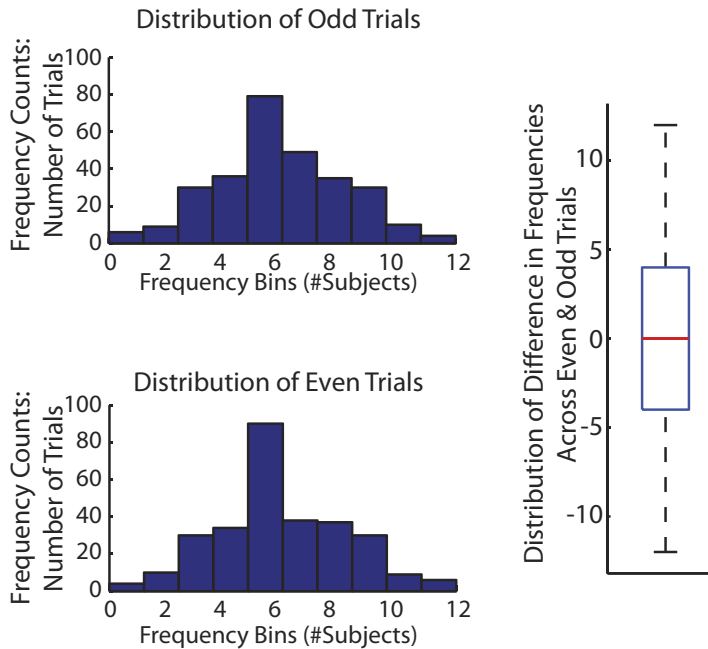
E-mail address: mahon@rcbi.rochester.edu (B. Z. Mahon).

Abbreviations: FDR, false discovery rate; MVPA, multi-voxel pattern analysis; ROIs, region of interest.

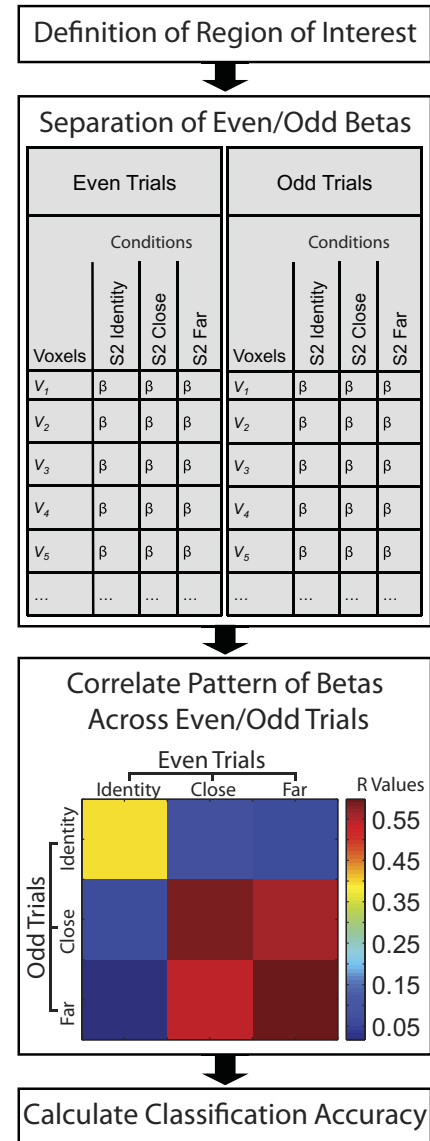
A. Trial Structure and Participants' Task



B. Split-half Analysis (Even/Odd trials) of Experimental Design



C. Schematic of Pattern Analysis



**Fig. 1.** (A) Structure of trials and analysis. Each stimulus (S1 and S2) consisted of an auditorally spoken word. Participants indicated the conceptual similarity between the pair of words by means of a button response after hearing a response cue (200 ms tone). (B) Functional data were divided into two halves, based on even and odd trials. Because the experimental design was distributed randomly with constraints across trials, the distributional analyses show that the design was balanced across the two halves of the data (even and odd). There were 288 unique trials per subject, and 12 subjects. Thus, the histograms indicate that the distribution of those 288 unique trials was distributed evenly across the even and odd trials, across all subjects. The boxplot substantiates this by representing the difference scores between the two distributions. (C) This panel shows schematically the analysis strategy: Regions of Interest (ROIs) were defined by contrasting S2 against S1, using only even trials for S2. We then tested for differences, using conventional averaging approaches, among Identity, Close, and Far for S2, for odd trials only. We also computed within- and between-condition correlations, always comparing even and odd trials. Classification accuracy was then calculated on the basis of the resulting correlation matrices.

the direct confrontation between the “close” and “far” conditions: one expectation on the basis of previous research (e.g., Rips et al., 1973) is that more fine-grained analysis is required in order to judge the similarity between very close concepts. Associated with such fine-grained processing are increased response times, and it may be predicted, differential BOLD responses, at least in those regions that are sensitive to the degree to which the concepts must be analyzed according to the task.

**EXPERIMENTAL PROCEDURES**

**Participants**

Fifteen healthy native Italian speakers, with normal eyesight (corrected with MR compatible goggles where necessary) and normal hearing participated in the experiment. The datasets for three participants were discarded due to excessive head motion. Of the 12 subjects included in the analysis seven were female, 10 were right handed (assessed with Edinburgh Handedness Inventory, Oldfield, 1971), and the group had a mean age of 32.4 years

(range: 20–51 yrs). Informed written consent was obtained under approved University of Trento and Harvard University protocols for the use of human participants in research, in accordance with the Declaration of Helsinki.

### Stimuli and design

Ninety-six words referring to common object concepts were selected. For each word a digitally recorded wave file (22.050 kHz, 16 Bit, native Italian speaker, female) was prepared. Custom software (ASF, J. Schwarzbach) written in Matlab utilizing the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) was used for stimulus presentation.

MR data were collected at the Center for Mind/Brain Sciences, University of Trento, on a Bruker BioSpin MedSpec 4T. The scanning session consisted of two phases. In Phase I a high resolution ( $1 \times 1 \times 1$  mm<sup>3</sup>) T1-weighted 3D MPRAGE anatomical sequence was performed (sagittal slice orientation, centric Phase Encoding, image matrix=256×224 (Read×Phase), FoV=256×224 mm<sup>2</sup> (Read×Phase), 176 partitions with 1 mm thickness, GRAPPA acquisition with acceleration factor=2, duration=5.36 min, TR=2700, TE=4.18, TI=1020 ms, 7° flip angle). Phase II of the scanning session consisted of the principal experimental task. Functional data were collected using an echo planar 2D imaging sequence with phase over-sampling (Image matrix: 70×64, TR: 2250 ms TE: 33 ms, Flip angle: 76°, Slice thickness=3 mm, gap=.45 mm, with 3×3 in plane resolution). Volumes were acquired in the axial plane in 37 slices. Slice acquisition order was ascending interleaved odd-even.

### Procedure

On each trial within Phase II, participants heard an auditorily spoken word (S1) followed at a jittered interval (2 to 8 s, in steps of .5 s, distribution with hyperbolic density) by a second auditorily presented word (S2). Participants were instructed to think about the conceptual similarity of the two objects referred to by S1 and S2. A jittered interval (2 to 8 s) after the onset of S2, an auditory tone (duration 200 ms) signaled for participants to enter their judgment of conceptual similarity by means of a button response, on a scale of '1 to 4' (two buttons for the right hand, two for the left hand). The next trial (i.e., onset of S1) began at a jittered interval (4 to 10 s) after presentation of the auditory response cue. The two auditorily presented words on each trial could be identical, semantically close, or relatively semantically far. The two stimuli always came from the same semantic category (animal, fruit/vegetable, tool, or non-manipulable nonliving). There were 24 items within each of the four semantic categories, by three conditions (Identity, Close, Far), resulting in 288 trials per subject (576 auditorily presented words). As we were not interested in modulation of BOLD responses as a function of semantic category, all analyses reported herein collapse across the semantic category of the stimuli. Each of the 96 auditorily spoken words was presented six times, three times as an S1 and three times as an S2. The experiment was divided into three runs of 96 trials. Stimulus order was randomized differently for each subject, with the constraints that (1) every 12 trials each level of conceptual similarity was presented four times, and (2) each semantic category was presented three times. In this way, it was ensured that all aspects of the experimental design appeared with equal probability throughout the scanning session. Within every run, all words appeared exactly twice.

All MR data were analyzed using Brain Voyager (v. 1.9). The first two volumes of functional data from each run were discarded prior to analysis. Preprocessing included, in the following order, slice time correction (sinc interpolation), motion correction with respect to the first volume in the run, and linear trend removal in the temporal domain (cutoff: three cycles within the run). Functional data were registered (after contrast

inversion of the first volume) to high-resolution de-skulled anatomy on a participant-by-participant basis in native space. For each individual participant, echo-planar and anatomical volumes were transformed into the standardized Talairach and Tournoux (1988) space. The data were not spatially smoothed in order to preserve spatial information about the distribution of BOLD responses at the voxel-level. All activation maps are projected onto the inflated anatomy of a single participant normalized to Talairach space. All functional data were analyzed using the general linear model. Events were convolved with a standard dual gamma hemodynamic response function. Beta estimates were standardized (z-scores) with respect to the entire time course. In addition to the regressors for the experimental conditions, six regressors of no interest were included, corresponding to the first derivative of the motion parameters obtained during preprocessing.

### Multi-voxel pattern analysis

Software written in Matlab, utilizing the BVQX Matlab toolbox (by Jochen Weber: [http://wiki.brainvoyager.net/BVQX\\_Matlab\\_tools](http://wiki.brainvoyager.net/BVQX_Matlab_tools)) was used to calculate within- and between-condition correlations, as well as classification accuracy. The data were split in half, according to even and odd trials across the three runs. In other words, the first, third, fifth (and so on) trials were classified as odd, while the second, fourth, sixth (and so on) trials were classified as even. Because stimulus order was random with the above described constraints, the experimental materials were distributed evenly across subjects across the two halves of the data. Nevertheless, in order to ensure that the experimental materials were evenly distributed across the two halves of the data, we analyzed the distribution of the unique trials ( $n=288$ ) across the two halves (even and odd) of the data. The results, plotted as histograms in Fig. 1B, demonstrate that the experimental design was evenly distributed across the two halves of the data according to their order of presentation (even/odd). This is substantiated statistically by the fact that the two distributions do not differ ( $t < 1$ ); the boxplot in Fig. 1B that plots difference scores between the two distributions has a median at  $\sim$ zero.

Within each Region of Interest (ROI) a matrix of *voxel* rows and *condition* columns (Identity, Close, Far) was extracted for S2s, separately for the odd and even halves of the data. The vector of beta values for a given condition across the voxels was then correlated within that condition (always correlating the "even" dataset with the "odd" dataset) and between that condition and the other conditions (again, "even" with "odd"). The analysis of classification accuracy followed Haxby and colleagues (2001) and Spiridon and Kanwisher (2002). Randomly selected groups of  $n$  number of voxels from the ROI were selected, with  $n$  ranging from 10 to 90, with an additional analysis over all voxels in the region. The number of times that  $n$  number of voxels were sampled was proportional to  $n$  (see Spiridon and Kanwisher, 2002). For ROIs for which there were not 100 voxels but rather  $x$  number (e.g., 49) of voxels (Medial Prefrontal, and Early Visual ROIs), randomly selected groups of  $n$  number voxels, with  $n$  ranging from  $x/10$  to  $x$  in steps of  $x/10$  were selected. Classification was considered accurate (i.e., 100%) on each iteration if the means of the within condition correlations for two conditions was greater than the between condition correlations. The results of this analysis are plotted in Fig. 3B (data points and error reflect the means and standard error of the means across participants for classification accuracy, calculated separately for each participant). The same procedure was followed to simulate chance (dotted lines, Fig. 3B), except that the beta values from the same ROI were randomized across voxels before calculating classification accuracy.

To statistically analyze classification accuracy,  $t$ -tests (two-tailed) were carried out at each sampling interval. For instance, a  $t$ -test was carried out comparing classification accuracy for Iden-

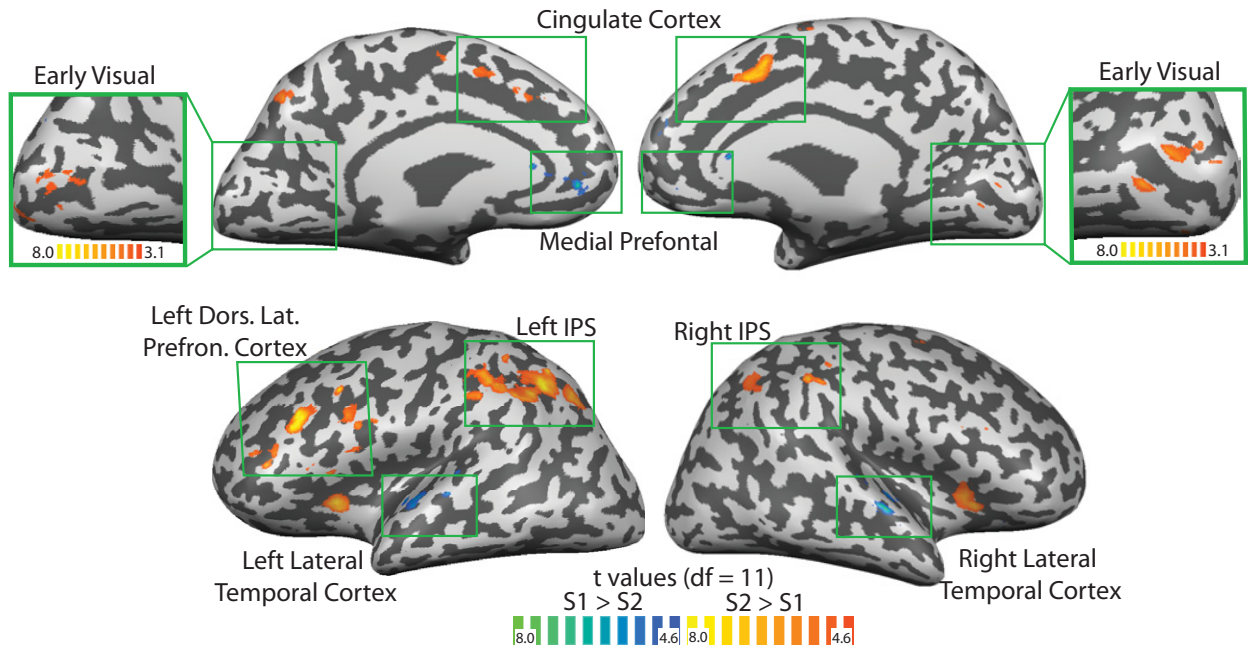
tivity vs. Far (real data) against Identity vs. Far (simulated chance), with 11 degrees of freedom (12 subjects). Thus, for every test of classification accuracy (Identity vs. Far, Identity vs. Close, and Close vs. Far) ten *t*-tests were carried out, one for every *n* number of voxels that were sampled. Fig. 4 plots the results of those *t*-tests for classification accuracy (each boxplot represents the distribution for all ten *t*-tests carried out for the conditions). Also plotted in Fig. 4 (green dots) are the *t*-values (absolute values of the *t*

statistic) obtained by carrying out conventional ROI contrasts comparing the same conditions (Identity vs. Far, Identity vs. Close, and Close vs. Far).

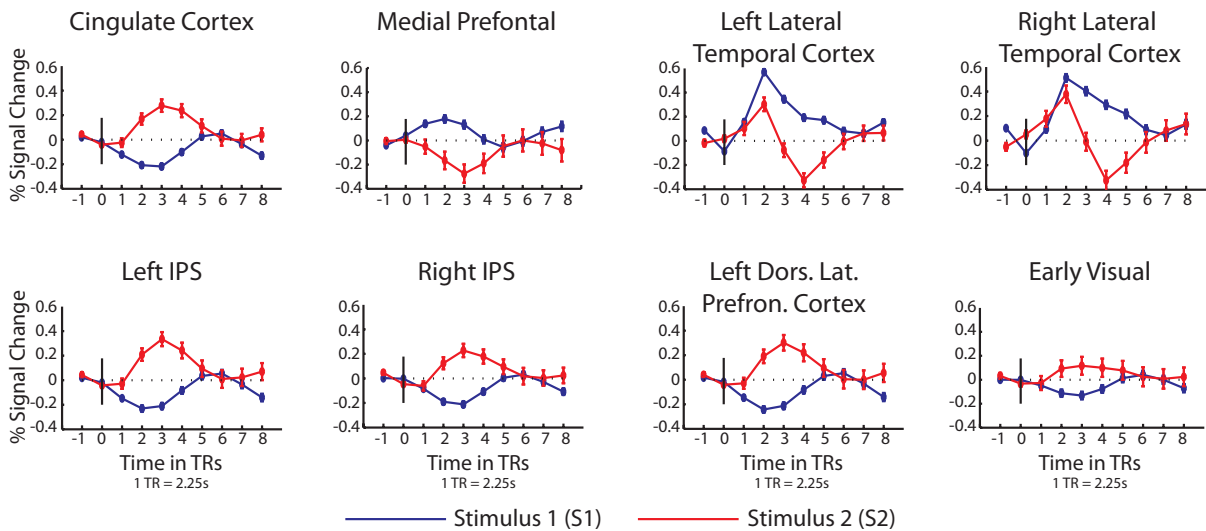
**Spotlight analysis of classification accuracy**

In the “spotlight” approach to studying classification accuracy at the whole brain level (Kriegeskorte et al., 2006), classification

**A. Definition of Regions of Interest: Stimulus 2<sub>(even)</sub> vs. Stimulus 1**



**B. Time Courses from Regions of Interest for S2<sub>(odd)</sub> and S1**

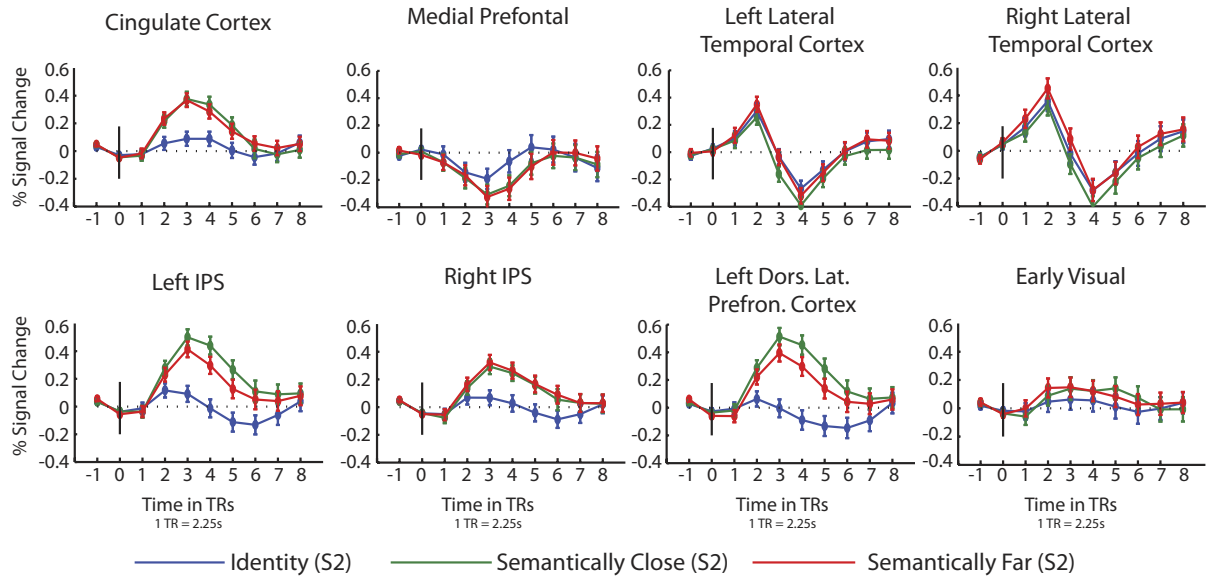


**Fig. 2.** (A) Regions of Interest (ROIs) were defined by contrasting the BOLD responses associated with processing the second stimuli of the pairs (S2s) against the BOLD responses associated with processing the first stimuli of the pairs (S1s) ( $P(\text{FDR}) < .05$ , whole brain). Areas identified as ROIs are labeled in the Figure and were then studied with further analyses. The early visual ROI was present at the  $P(\text{FDR}) < .05$  threshold but was defined for further analysis at the more relaxed threshold shown in the inset ( $P < .005$ ). This was to have a sufficient number of voxels to be able to carry out multi-voxel pattern analyses. (B) The event-related time courses for S1s and S2s were extracted from the ROIs. Importantly, the time courses for S2s that are plotted are independent of the definition of the ROIs, because the ROIs were defined using even trials for S2s, while the time courses were calculated over only odd trials for S2s.

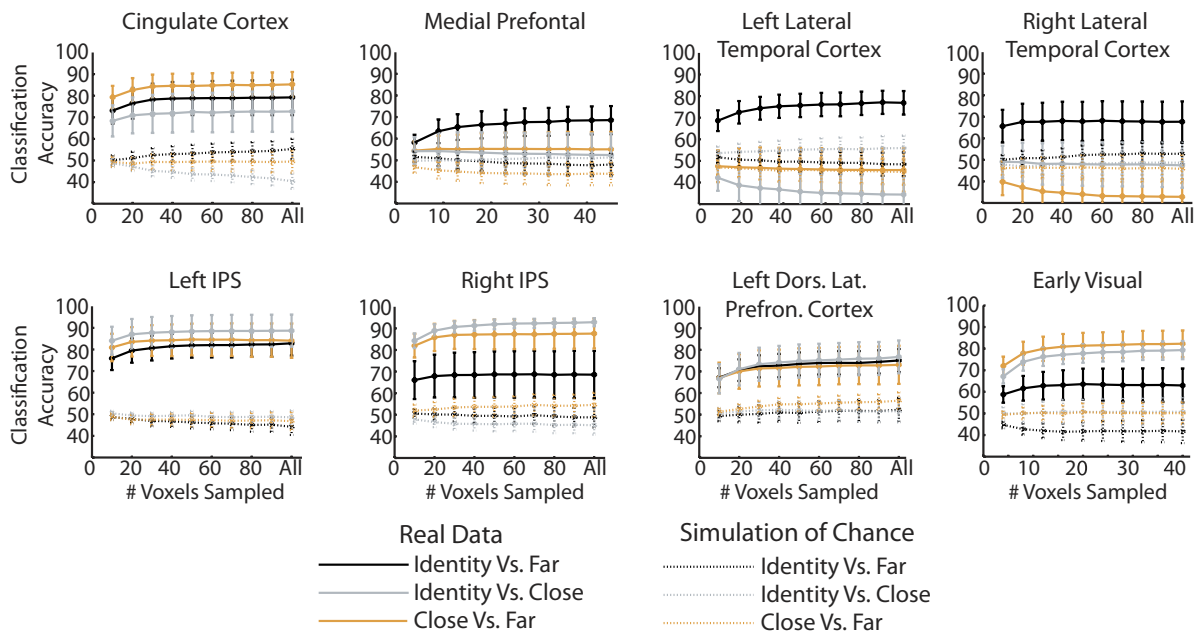
accuracy was mapped throughout the brain. The “spotlight” passes over each (*i*th) voxel (in each participant), and extracts the beta estimates (for even and odd presentations and for all experimental conditions) for the cube of voxels ( $n=27$ ) that surround the *i*th voxel. The correlation coefficients for within- and between-condition correlations were calculated across that set of 27 voxels and Fisher transformed to *t*-values. The difference in *t*-values between the within- and between-condition

correlations was then written to the *i*th voxel within a new array. The contrast map plots the results of a one-sample *t*-test (two tailed, degrees of freedom=11) computed within each voxel across participants on the difference in *t*-values for the within- vs. between-condition correlations (averaging over the three conditions). Thus, the contrast map in Fig. 5A plots voxels that show differential within- vs. between-condition correlations at the group-level.

A. Time Courses from Regions of Interest for S2<sub>(odd)</sub> according to Semantic Distance in relation to S1



B. Classification Accuracy for Multi-Voxel Pattern Analysis within Regions of Interest for S2<sub>(odd)</sub>



**Fig. 3.** (A) Event-related time courses were extracted from the ROIs for odd S2 trials, separately for the three experimental conditions (Identity, Close, Far). Error bars reflect standard errors of the means, across subjects. The x-axes are in units of TRs, with each TR equaling 2.25 seconds. (B) Classification accuracy (y axes) by experimental condition (solid lines) is plotted for randomly selected sets of voxels of *n* number (x axes) for each ROI (see methods for details). The same procedure was followed to simulate chance, except that the beta values were randomized across voxels and experimental conditions (dotted lines). Data points and error reflect the means and standard error of the means across participants for classification accuracy.

## RESULTS AND DISCUSSION

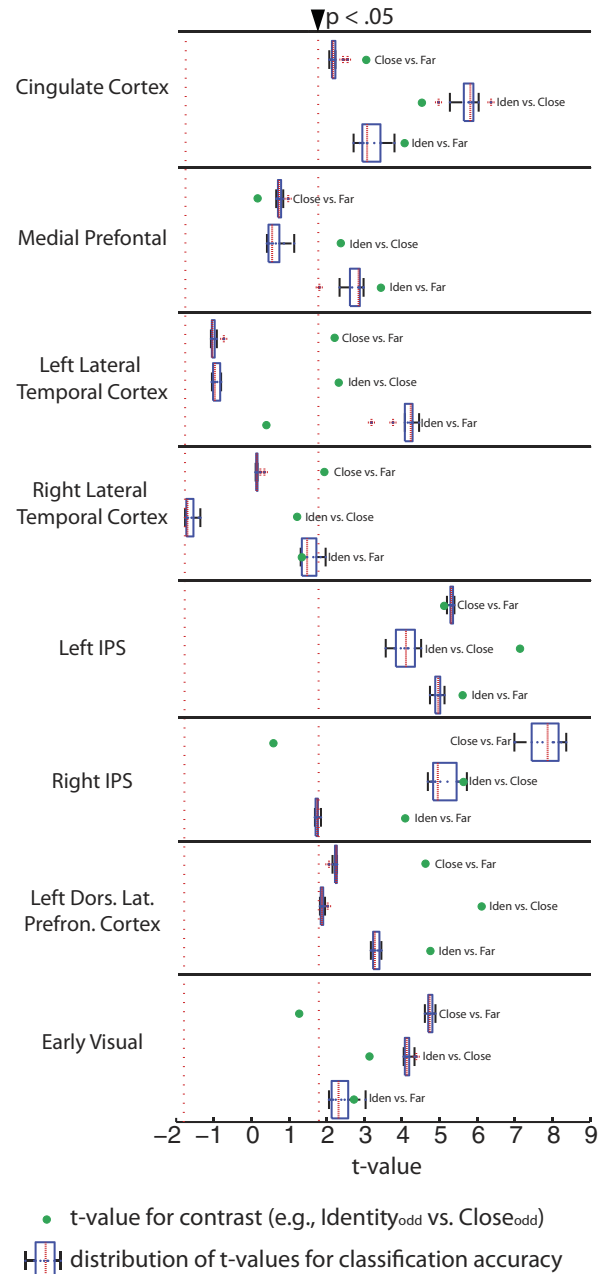
We first analyzed participants' behavioral responses, collected during scanning, to ensure that the manipulation of semantic distance in our materials was salient to participants. Participants' judgments reflected the manipulation of conceptual similarity (standardized ratings (z-scores)  $\pm$  sem: Identity =  $-1.10 \pm 0.03$ ; Close =  $0.07 \pm 0.05$ ; Far =  $1.03 \pm 0.03$ ;  $F_{2,22} = 512.9$ ,  $P < .001$ ; all pair-wise tests ( $t$ -tests, two tailed) significant at  $P < .001$ ).

In order to define ROIs we created a statistical contrast map that identified regions that were (1) more activated during processing of the second members of the pairs than during the first members of the pairs ( $S2 > S1$ ), as well as the reverse ( $S1 > S2$ ) (thresholded at  $P(\text{FDR}) < .05$ , that is, corrected for False Discovery Rate (FDR) for the whole brain volume). Critically, this contrast map was defined using S2 stimuli from even trials only. The use of "even" trial S2 stimuli for the definition of ROIs allowed us to then test the hypotheses of interest with separate data, provided by the "odd" trials (Kriegeskorte et al., 2009). Fig. 2 (Panel A) shows the resulting contrast map. Larger BOLD responses for S2 compared to S1 (red–yellow color scale) were observed in regions typically observed as part of the classic working memory and control system, including medial and lateral parietal and frontal regions (see e.g., Hickok and Poeppel, 2007; Pinel et al., 2004; Repovš and Baddeley, 2006; Smith and Jonides, 1999; Turken and Swick, 1999). In line with previous findings (for review see Hickok and Poeppel, 2007) there was a bias toward stronger effects in the left hemisphere, probably reflecting the language component of the task. However, while stronger in the left hemisphere, the effect in parietal cortex along the intraparietal sulcus was bilateral. In addition, there was a smaller volume of activation in early visual cortex, in the vicinity of V2/V3 (Dougherty et al., 2003). The increased BOLD activity in this early visual region is most naturally explained by the visual imagery demands associated with retrieving information about specific object concepts (Kosslyn et al., 2001). Potentially relevant in this regard is the previous observation that early "visual" regions track verbal working memory performance in blind subjects (e.g., Amedi et al., 2003).

Also shown in the contrast map of Fig. 2A are regions showing relatively stronger BOLD responses to the first stimuli of the pairs (blue–green color scale); those regions included medial prefrontal cortex and bilateral lateral temporal cortex. The relatively increased BOLD responses for S1 than S2 in medial prefrontal cortex overlaps with a component of the "resting" or "default" network, which typically shows tonically higher levels of activation during non-task periods, as well when cognizing about mental states (e.g., Gusnard et al., 2001; Mitchell et al., 2002). The fact that this region showed a larger BOLD response during processing of S1 than S2 is consistent with the demands of the task in that participants could actively compare the two concepts on a given trial only after having heard the second stimulus. In line with this interpretation, there were relatively larger BOLD responses (or less deactivation) in the same region for the relatively more demanding experimental

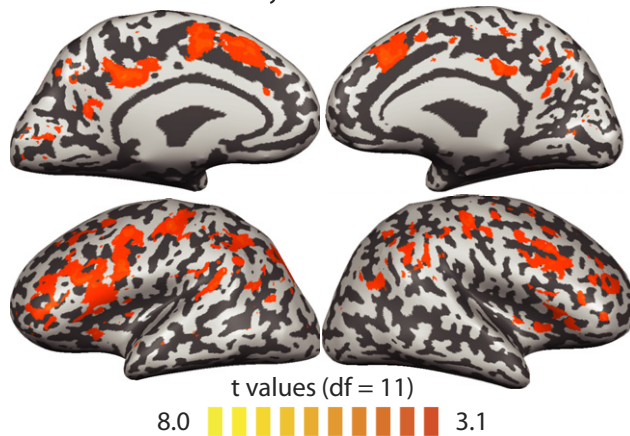
conditions ("close and far," compared to "identity"; see Fig. 3A).

BOLD responses were also elevated for S1 compared to S2 in lateral temporal regions in or near auditory cortex. This observation is consistent with two non-mutually ex-

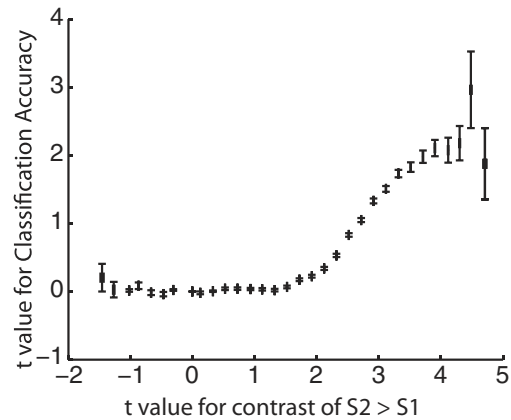


**Fig. 4.** Summary of all statistical analyses for conventional averaging (green dots) and classification accuracy (box plots). A series of planned  $t$ -tests were carried out within each ROI, averaging the beta values over all voxels within the ROI, and contrasting Identity to Far, Identity to Close, and Close to Far; this analysis was carried out over only S2s from odd trials. Each green dot represents the resulting  $t$ -value for the respective contrast and ROI (all  $t$ -values plotted as absolute values; see Fig. 3A for polarity of the effects). Boxplots within the same figure and on the same axis represent the distribution of  $t$ -values obtained by contrasting each solid line in Fig. 3A to its respective baseline (dotted line of same color in same graph).

### A. Whole-brain level spotlight analysis of classification accuracy



### B. Whole-brain level relation between classification accuracy and differential response to S2 vs. S1



**Fig. 5.** (A) Whole-Brain Analysis of Classification Accuracy. The group-level statistical overlay shows the results of a whole brain search for voxels that carry information about the semantic context in which participants were thinking about the pairs of object concepts. Significant effects in the contrast map reflect voxels that show differential within- vs. between-condition correlations at the group-level. (B) The graph plots, on a voxel-by-voxel basis, classification accuracy (y-axis; i.e.,  $t$ -values from Panel A) against the contrast weighted  $t$ -value from S2 vs. S1 (x-axis; i.e.,  $t$ -values from Fig. 2A; data binned in steps of .2 increments of  $t$ -values). The positive and monotonic relationship indicates that classification accuracy among the different experimental conditions of S2 was greater for voxels that also showed relatively stronger BOLD responses for S2 compared to S1.

clusive interpretations. The first is that, because participants had to hold the S1s in memory until hearing the S2s, the relatively larger and longer (i.e., higher in amplitude and sustained) BOLD responses in this region may reflect auditory memory for the stimulus (Hickok and Poeppel, 2007). However, that interpretation cannot be distinguished, on the basis of the current dataset, from the possibility that averaging of the time course will include within its umbrella S2 events, such that the response to S2 may be partially represented in the time courses for S1, therefore accounting for the sustained activity.

#### Analysis of conceptual similarity between S2 and S1

We used a combination of conventional ROI-based contrast analyses and multi-voxel pattern analysis (MVPA) techniques (see methods for details) to test whether the brain responses associated with processing the second members of the pairs are modulated as a function of the conceptual similarity between the two members of the pair. For example, is the brain response associated with thinking about the object concept “chair” modulated as a function of whether the previously presented word was “chair”, “stool”, or “stove”?

The event-related time courses illustrating the amplitude and profile of BOLD responses to S2s across the three levels of conceptual similarity for all ROIs is plotted in Fig. 3A. The  $t$ -values corresponding to the three possible contrasts (Identity vs. Far, Identity vs. Close, and Close vs. Far) are plotted in Fig. 4 (green dots). In a parallel analysis, we tested whether it was possible to distinguish Identity vs. Far, Identity vs. Close, and Close vs. Far on the basis of the distributions of beta values across the voxels within each of the ROIs. In other words, this analysis approach harnessed the variation in BOLD responses across voxels within the ROIs, rather than averaging across that variation. The results of the analysis of classification accuracy for all

ROIs is plotted in Fig. 3B. The corresponding  $t$ -values for classification accuracy are plotted in Fig. 4 (boxplots).

As can be seen in Fig. 4, there were three regions for which both conventional ROI-based contrasts and classification accuracy were significant for all three confrontations of the experimental conditions (Identity vs. Far, Identity vs. Close, and Close vs. Far): Left IPS, Left Dorsal Lateral Prefrontal Cortex, and the anterior Cingulate cortex. All three of these regions also showed (Fig. 2) differential responses to S2 compared to S1. The other ROIs that were identified as showing larger BOLD responses to S2 compared to S1 in Fig. 2, namely Right IPS and early visual cortex, did not show significant differences using conventional averaging approaches for the contrast of the Close and Far conditions. However, for the same comparison in those regions (Close vs. Far) classification accuracy was significantly above chance. Finally, in ROIs showing relatively greater BOLD responses to S1 compared to S2 (medial prefrontal cortex and lateral temporal cortex), classification accuracy tended to not be significantly different than chance; the clear exception was the contrast of Identity vs. Far in left lateral temporal cortex, for which classification accuracy was significant. That pattern is consistent with the possibility that phonological similarity may be a principal determinant of the distribution of BOLD responses in that region.

#### Mapping classification accuracy throughout the brain

We then studied classification accuracy in a whole brain analysis. Voxels throughout the whole brain were coded according to the degree to which they carried information about the conceptual similarity between the two object concepts in the pairs (Fig. 5; for previous studies using this approach, see Haynes et al., 2007; Kriegeskorte et al., 2006). This analysis highlighted the medial and lateral

parietal and frontal regions previously identified, as well as voxels in the vicinity of the early visual ROI. Additional regions were also identified, according to the role that those regions played in the task in which participants were engaged. For instance, because participants were entering their judgments with button responses, high classification accuracy was observed in motor cortex.

One pattern that emerged in this analysis is that regions that tended to show relatively stronger responses for S2s compared to S1s (red–yellow color space in Fig. 2A) also tended to show relatively high classification accuracy (see also Fig. 4). To confirm that this was in fact the case, we correlated classification accuracy (*t*-values) throughout the whole brain, voxel-by-voxel, with the *t*-values for the contrast of S2 vs. S1. If, as would be expected from the findings described above, regions showing increased responses to S2 also show relatively higher classification accuracy, there should be a positive monotonic relationship between the two measures. The results of the analysis, plotted in Fig. 5B, support this expectation. This relationship may be due to the fact that regions showing overall higher BOLD responses to S2 have greater variability across the population of voxels, which is harnessed by the analysis of classification accuracy.

## CONCLUSION

The cognitive situation presented by our experiment was designed to be a controlled version of a process that occurs commonly in normal life: we think about multiple objects and make decisions based on knowledge that we have about those objects. Our findings indicate that the semantic context within which we think about object concepts determines a spatial distribution of brain responses in regions of the brain that are critical for representing and processing object properties. An issue for future research is whether such sub-threshold “scattering” of BOLD responses across a set of voxels is driven by the presence of sub-regions (i.e., voxels) within the broader set that express relatively high functional connectivity to regions that are principally involved in the task. More generally, and independently of the exact cause of subthreshold, distributed information across voxels, our findings indicate that there is specificity of brain responses not only at the level of the content of what we think about, but also according to the broader context in which we have those thoughts.

*Acknowledgments*—The authors are grateful to Jorge Almeida, Jessica Cantlon, Angelika Lingnau, Jens Schwarzbach, Kyle Simmons, and Alex Martin for discussing the research. BZM was supported in part by an Eliot Dissertation completion grant and NIH training grant 5 T32 19942-13; AC was supported by grant DC006842 from the National Institute on Deafness and Other Communication Disorders. Preparation of this article was supported by a grant from the Fondazione Cassa di Risparmio di

Trento e Rovereto. Additional support was provided by Norman and Arlene Leenhouts. BZM is now at the Department of Brain and Cognitive Sciences, University of Rochester.

## REFERENCES

- Amedi A, Raz N, Pianka P, Malach R, Zohary E (2003) Early “visual” cortex activation correlates with superior verbal memory performance in the blind. *Nat Neurosci* 6:758–766.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Cantlon JF, Platt LM, Brannon EM (2009) Beyond the number domain. *Trends Cognit Sci* 13:83–91.
- Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J (2003) Visual field representations and locations of visual areas  $V_{2d}$  3 in human visual cortex. *J Vis* 3:586–598.
- Gusnard DA, Akbudak E, Shulman GL, Raichle ME (2001) Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci U S A* 98:4259–4264.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
- Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RE (2007) Reading hidden intentions in the human brain. *Curr Biol* 17:323–328.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8:393–402.
- Kosslyn SM, Ganis G, Thompson WL (2001) Neural foundations of mental imagery. *Nat Rev Neurosci* 2:635–652.
- Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI (2009) Circular analysis in systems neuroscience—the dangers of double dipping. *Nat Neurosci* 12:535–540.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Martin A (2007) The representation of object concepts in the brain. *Annu Rev Psychol* 58:25–45.
- Mitchell JP, Heatherton TF, Macrae CN (2002) Distinct neural systems subserve person and object knowledge. *Proc Natl Acad Sci U S A* 99:15238–15243.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Pinel P, Piazza M, Bihan DL, Dehaene S (2004) Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* 41:1–20.
- Repovš G, Baddeley A (2006) The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience* 139:5–21.
- Rips LJ, Shoben EJ, Smith EE (1973) Semantic distance and the verification of semantic relations. *J Verbal Learn Verbal Behav* 12:1–20.
- Smith EE, Jonides J (1999) Storage and executive processes in the frontal lobes. *Science* 283:1657–1661.
- Spiridon M, Kanwisher N (2002) How distributed is visual category information in human occipital-temporal cortex? An fMRI study. *Neuron* 35:1157–1165.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain New York, NY: Thieme.
- Turken AU, Swick D (1999) Response selection in the human anterior cingulate cortex. *Nat Neurosci* 2:920–924.