

Neural correlates of partial lexical activation

Kathleen Pirog Reville*^{†‡}, Richard N. Aslin*, Michael K. Tanenhaus*, and Daphne Bavelier*

*Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627; and [†]School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332

Communicated by Elissa L. Newport, University of Rochester, Rochester, NY, July 21, 2008 (received for review December 1, 2007)

As a spoken word unfolds over time, it is temporarily consistent with the acoustic forms of multiple words. Previous behavioral research has shown that, in the face of temporary ambiguity about how a word will end, multiple candidate words are briefly activated. Here, we provide neural imaging evidence that lexical candidates only temporarily consistent with the input activate perceptually based semantic representations. An artificial lexicon and novel visual environment were used to target human MT/V5 and an area anterior to it which have been shown to be recruited during the reading of motion words. Participants learned words that referred to novel objects and to motion or color/texture changes that the objects underwent. The lexical items corresponding to the change events were organized into phonologically similar pairs differing only in the final syllable. Upon hearing spoken scene descriptions in a posttraining verification task, participants showed greater activation in the left hemisphere anterior extent of MT/V5 when motion words were heard than when nonmotion words were heard. Importantly, when a nonmotion word was heard, the level of activation in the anterior extent of MT/V5 was modulated by whether there was a phonologically related competitor that was a motion word rather than another nonmotion word. These results provide evidence of activation of a perceptual brain region in response to the semantics of a word while lexical competition is in process and before the word is fully recognized.

fMRI | MT/V5 | semantic representations | spoken word recognition

Spoken words unfold over time as a series of rapidly changing acoustic events. To comprehend spoken language, the listener must map this speech signal onto representations of meaning. Behavioral (1–3) and neuroimaging (4) evidence suggests that multiple words that are temporarily consistent with the unfolding acoustic input compete for recognition. Although experiments using a variety of behavioral paradigms have shown that syntactic, semantic, and nonlinguistic features of the words in the active competitor set are available before word recognition is complete (5–9), there is no neural evidence indicating whether semantic representations are involved in this competition. However, other behavioral tasks have shown interactions between perceptual processing and language processing (10–13), and recent neuroimaging evidence suggests that the concepts that are activated by written or spoken words are distributed across cortical regions intrinsically involved in perception and action (7, 14–22). Nevertheless, there is debate about whether the recruitment of these semantically induced activations in perceptual brain areas is automatic or strategic (23). In this article, we demonstrate that even lexical candidates that are only temporarily consistent with the unfolding word activate brain regions involved in perceptual processing, thus providing clear evidence that lexical processing automatically activates perceptually based semantic representations.

Finding a neural correlate of semantic activation during spoken word recognition requires a novel approach, because it is extremely challenging, if not impossible, to find words with the appropriate phonological and semantic characteristics in English or any other existing language because form and meaning are largely unrelated in natural languages. For example, the words *candle* and *candy* have similar phonological forms that differ only

at the final syllable, but their meanings are quite different. Here, we controlled the relationship between the acoustic/phonetic overlap of pairs of words (cohort competitors) and their semantic properties by using an artificial lexicon. Previous behavioral research has established that artificial lexicons show the same signature effects observed during the processing of real language (see ref. 2 for an extended discussion). It is also clear from several training studies that links between novel semantic features and perceptual brain areas can be acquired in a short-term experimental context (24–26). We therefore trained participants to associate novel spoken word forms with objects and events with controlled perceptual properties, allowing us to focus on finding evidence for lexical activation of semantic representations within a region of interest (ROI) previously found to be activated by the semantic properties of words, human MT-MST/V5 (hereafter MT/V5). MT/V5 responds strongly to visual motion, and its anterior extent has been associated with the processing of motion-related words (15–17, 19–21, 27).

Over several days of training, participants learned to map new lexical items to a set of novel visual objects with two kinds of perceptually based semantic properties, relating either to changes in motion trajectory or surface appearance. The degree of semantic similarity between the members of phonologically related competitor sets was systematically varied so that the items in some competitor sets had similar perceptual correlates, whereas in other sets each item had distinct perceptual correlates. We focused on MT/V5 and its anterior extent as the ROI associated with the perception of one type of change present in our stimuli, visual motion, and examined how activation differed when participants heard a newly learned word as a function of the semantic features of that word and the semantic features of its phonological competitor. Our primary goal was to obtain a neural measure of partial lexical activation by cohort competitors.

A secondary goal was to more precisely determine where in the greater MT/V5 complex semantic properties of a lexical item lead to robust activation. Previous studies have shown that MT/V5 responds strongly to imagined or implied visual motion (28, 29) in addition to direct visual stimulation (30–33). MT/V5 is also active when participants perform a similarity judgment about pictures depicting actions (15, 16). Several studies have suggested that this word-responsive area is in fact the same as the perceptually driven area (34, 35). However, studies that have functionally localized MT/V5 have found that the word-responsive area is just anterior to left MT/V5 in the middle temporal gyrus (15, 16). Given the varied experience that humans have with words and with real-world objects and events, it is impossible to know whether the referent of a printed word contacts the same semantic knowledge in all participants and how much overlap there is between the representation activated by reading the word and the representation activated by viewing a drawing the experimenter uses to convey the same concept. In

Author contributions: K.P.R., R.N.A., M.K.T., and D.B. designed research; K.P.R. performed research; K.P.R., R.N.A., M.K.T., and D.B. analyzed data; and K.P.R., R.N.A., M.K.T., and D.B. wrote the paper.

The authors declare no conflict of interest.

[†]To whom correspondence should be addressed. E-mail: kate.pirog@psych.gatech.edu.

© 2008 by The National Academy of Sciences of the USA

Table 1. Sample stimuli

Condition	Example
Motion target	biduko goki (horizontal oscillation)
Motion cohort	biduka goki (rotation)
Motion target	dukobi kitu (size oscillation)
Nonmotion cohort	dukoba kitu (speckling)
Nonmotion target	gapito ripa (whitening)
Motion cohort	gapitu ripa (vertical oscillation)
Nonmotion target	potagu tumi (marbling)
Nonmotion cohort	potagi tumi (blackening)

addition to controlling the phonetic, semantic, and perceptual properties of the stimuli, the use of an artificial lexicon allowed us to carefully control the amount and types of experience participants had with the stimuli. In the present study, the only possible referent for the newly learned lexical item was the one provided during training.

Results

Behavioral Training. During 3 days of behavioral training (see *Methods* for details about the training procedure), participants learned that eight novel bisyllabic CVCV lexical items referred to eight novel shapes, and that eight trisyllabic CVCVCV items referred to eight changes that the shapes could undergo. Four of these changes caused a shape to move (horizontal oscillation, vertical oscillation, expansion-contraction oscillation, or rotation) and four caused changes to the shape's surface appearance (blackening, whitening, marbling, or speckling). All of the novel shape names were phonologically unique, but the items referring to the change events were arranged into four pairs of onset-overlapping "cohort" competitors differing only at the last vowel (Table 1). Both members of one cohort pair referred to motion changes, and both members of another pair referred to surface changes, but for the remaining two cohort pairs, one member of each pair referred to a motion change and the other to a surface change. Participants were not explicitly informed about the relationship between the phonological forms and the meanings of the referents, nor were they told that the motion changes and surface changes formed categories or sets of any sort. By the end of the third day of training, all participants had reached ceiling levels of performance (> 98% correct) on a match/no-match task in which they indicated whether a spoken two-word phrase (consisting of the name of a change event and the name of an object) correctly described a scene containing a changing object that was then presented on the computer screen. On the fourth day of the experiment, participants performed a modified version of this task during imaging. Although performance did decrease slightly between behavioral and fMRI sessions, participants still performed highly accurately (95% correct, range 93–99%).

Brain Imaging. Participants completed five runs of the experimental task. Each trial consisted of three major intervals: a spoken-word interval, a visual interval, and a response interval (Fig. 1). During the spoken-word interval, which lasted between 3 and 7 s, participants heard a two-word phrase referring to a change and an object. This phrase lasted for approximately the first 1,600 ms of the interval; the rest of the interval was silent. During the visual interval, one of eight objects undergoing one of eight changes was displayed on the screen for 2–4 s. During the following 3-s response interval, participants decided whether the stimuli from the spoken-word interval and the visual interval matched, indicated their decision with a button box, and received auditory feedback. Trials were separated by an intertrial interval (ITI) ranging from 1 to 9 s. Our analyses focus on an *a priori*

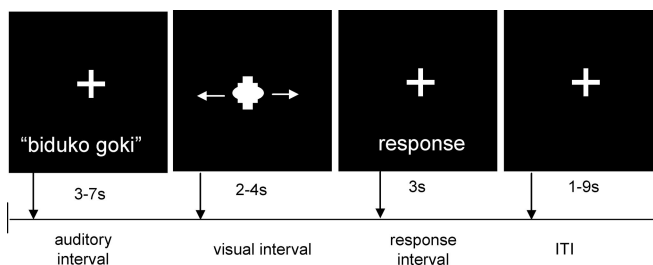


Fig. 1. Sample test trial timeline. Participants first heard the name of a change/object combination. Stimulus duration on all trials was $\approx 1,600$ ms, but the time between auditory stimulus onset and visual stimulus onset varied between 3 and 7 s. Visual stimuli were displayed for the entire visual interval of 2–4 s. After the visual stimulus disappeared, participants had 3 s to respond via button box and were given auditory feedback.

region of interest, MT/V5 and its anterior extent, previously shown to be recruited by the reading of motion words. After completing the experimental runs, participants were presented with 20-s alternating blocks of static and radially moving random dots to uniquely identify MT/V5. During analysis of the data from the experimental runs, all three event types (auditory interval, visual interval, and response interval) were included in the model. The outputs of the individual deconvolution analyses were normalized to a standard MNI brain for group analysis. Separate one-way mixed-effects ANOVAs were performed on the conditions of interest; for consistency, all analyses were performed at a threshold of $P = 0.01$, uncorrected.

Response to Visual Stimulus Presentation. Data from the visual interval allowed us to confirm that the presentation of the visual stimuli activates the desired perceptual brain areas. MT/V5 is active to a greater extent when participants view the motion-change events than when they view surface-change events. A conjunction analysis was performed to discover the area of activation common to viewing moving versus static dots and to viewing motion-change events versus surface-change events during the visual interval. A large area of overlapping activation was found in both hemispheres (MNI center of mass coordinates and cluster sizes: RH 43, – 67, 6, 199 voxels; LH – 41, – 71, 3, 131 voxels). This analysis confirmed that viewing the motion-change events recruits MT/V5 as does viewing moving dots, with the center of the activation located well within the standard boundaries of MT/V5 (31–33).

Response to Auditory Stimulus Presentation. We next examined whether hearing lexical items that refer to events with significant motion features activates MT/V5 and/or the cortex anterior to it more strongly than hearing lexical items referring to events without motion features. During the spoken-word interval of each trial, the participants heard the names of a change and an object but their visual referents had not yet been displayed. Any activation seen during this interval, therefore, can only be attributed to the analysis of the spoken words and not to bottom-up visual information. A cluster of voxels within the left hemisphere ROI is significantly more active when participants hear items referring to motion changes than when they hear items referring to texture or color changes (Fig. 2). Conjunction analysis reveals that 5 of the 27 active voxels for the auditory-interval contrast overlap with voxels from the visual-interval contrast at a threshold of $P = 0.01$ for both contrasts. Consistent with previous reports of activation anterior to MT/V5 for words, the center of mass of this 27-voxel cluster responding selectively to spoken motion words (– 42, – 59, 10) is anterior to the center of mass of the activation seen while participants are viewing the motion events (– 41, – 71, 3).

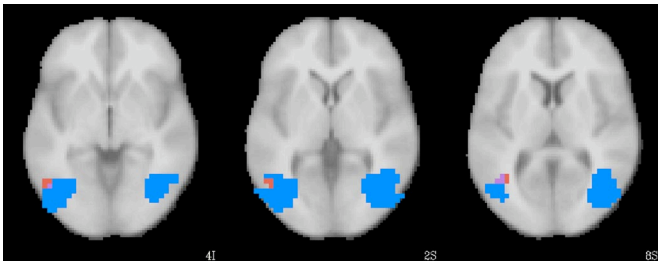


Fig. 2. Overlapping activation (purple) of the visual motion contrast (blue) and the auditory motion/non motion contrast (red) thresholded at $P = 0.01$, uncorrected. Note that the contrast for spoken words is left lateralized and overlaps with the anterior edge of the visual motion contrast. In all anatomical images, the left hemisphere is shown on the left. Functional data are overlaid on the MNI 152 average brain.

Having identified an area that is more responsive to spoken words referring to motion events than to texture and color events, we then asked whether the activation levels of voxels in this region may be modulated by “evidence for motion” from phonological cohorts. Although the unfolding speech is eventually consistent with only one word, initially the input is consistent with zero, one, or two motion words. We predict minimal effect of the temporarily consistent cohort competitor when the target word is a motion event because this target itself leads to strong activation. In contrast, we predict that the most sensitivity to the cohort competitor will be seen when the target does not refer to a motion event, as the background level of activity caused by the target word is low and small differences in activation caused by the semantics of the cohort may be visible. To address these predictions, a contrast analysis was performed on the voxels showing significant differences in activation for motion and nonmotion target words. Seventy-four percent (20 of 27) of the voxels in this region (Fig. 3A) were significantly correlated with the contrast using weights of +1, +1, -0.5, and -1.5 for motion target/motion cohort, motion target/nonmotion cohort, nonmotion target/motion cohort, and nonmotion target/nonmotion cohort conditions, respectively. Planned pairwise comparisons over the voxels showing the contrast trend were performed to confirm that the semantics of the cohort competitor affected activation levels when the target was a nonmotion word (Fig. 3B). As expected, when the target referred to a motion event, the semantics of the cohort did not significantly affect activation levels ($|t| < 1$). For nonmotion targets, however, we confirmed

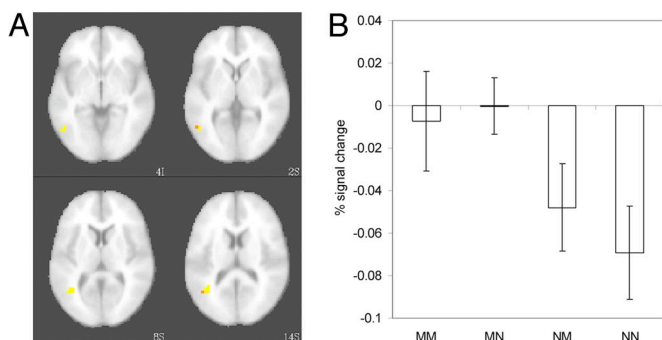


Fig. 3. The effect of target and competitor semantics on activation in the anterior extent of MT/V5. (A) Brain areas showing greater activity for words referring to motion events than texture events (yellow), with a subset of voxels (orange) significantly active for the contrast estimating activation of the phonological competitor at $P = 0.01$. (B) Average percent signal change from the voxels in the auditory region that contributed to the significant trend for each combination of target and competitor semantics for all participants. Error bars indicate between-subjects SEM.

our prediction that a cohort competitor referring to a motion event would result in higher activation levels than a cohort competitor that referred to a color- or texture-change event, ($t = 2.01$, $P < 0.05$, one-tailed). This finding indicates that the semantics of the cohort competitor do affect the overall activation pattern generated by temporarily ambiguous target words.

Given that competitor activation is not generally considered a conscious process, it is unlikely that this graded activation is a strategic or postdecision effect. Furthermore, on 33% of the no-match trials (20% of all trials), the visual event that participants saw actually contained the object change that was the correct referent of the cohort competitor. If participants misheard the word or were unsure of its meaning, they would have had high false-alarm rates on these trials. The data reported here come only from participants who performed accurately on these “catch” trials in addition to maintaining high overall performance levels (averaging 95% correct on all trials and 90% correct on catch trials). Therefore, it is unlikely that the responsiveness of activation levels to semantic features of both members of the competitor set is caused by lingering uncertainty about the identity of the target word past the point of disambiguation.

Discussion

By combining a well controlled artificial lexicon with a ROI whose level of activation is modulated by the perceptually based semantics of newly learned words, the results of the present experiment confirm the presence of partial activation before the point of disambiguation in spoken-word recognition. These results provide a neural imaging correlate of semantic activation of lexical competitors that previously was only available by using behavioral methods.

We find that both the target and the cohort motion stimuli activate a language-sensitive ROI that is left lateralized and anterior to the focus of activation for visually displayed moving versus static stimuli. Crucially, more activation was observed when a nonmotion target had a motion cohort than when it had a nonmotion cohort. We believe that these results reflect transient activation of the perceptually linked semantic representation of the cohort competitor during lexical competition and that we can rule out the postlexical conceptual processing or strategic visual imagery criticisms that other studies of semantic knowledge are vulnerable to. First, studies of spoken-word recognition using the visual world paradigm (36) show that participants frequently make eye movements to the competitor objects before eventually fixating and grasping the target object. These studies, in which participants are asked to pick up an object in the presence of other objects that are phonological competitors, show that the rise and fall of cohort competition are closely time-locked to the unfolding speech signal and are not strategic. Participants are generally unaware that they ever fixate the cohort competitors, but overall cohort competitor fixation proportions are reliably (although temporarily) high in studies with real words and in studies with artificial lexicons (2, 37–39), including a lexicon similar to the one used here (39). Second, it is clear that participants are not simply accepting a visual stimulus consistent with either final syllable of the ambiguous word or confusing the nonmotion cohort with the spoken target word, because they correctly reject the cohort’s visual referent after hearing the target. Furthermore, the activation levels within the ROI for the motion-target, nonmotion cohort condition and the nonmotion target, motion cohort condition are statistically different. If participants were strategically activating the meanings of both the target and the cohort competitor, the strategy needed to explain the observed activation patterns is quite complex because the contents of those pairs of representations are identical (one motion word, one nonmotion word). Finally, when explicit visual imagery is required of experimental participants, several studies have shown the resulting activation

Table 2. Artificial lexicon

Event words	Object words
/bi'duko/	/bado/
/bi'duka/	/dora/
/du'kobi/	/goki/
/du'koba/	/kitu/
/po'tagi/	/masi/
/po'tagu/	/pabu/
/ga'pitu/	/ripa/
/ga'pito/	/tumi/

patterns to be bilateral (29, 40) as opposed to the strongly lateralized pattern seen here.

In summary, the pattern of activation for newly learned motion words anterior to MT/V5 and the modulation of these activation levels caused by the semantic features of the competitor set suggest a very tight coupling between the neural substrate for lexical representations and perceptual brain regions. In addition, the techniques used here have allowed us to examine a question normally thought to depend on precise time-course measures to reveal modulations of activation in spatially relevant locations. Similar logic in using a ROI to investigate other short-lived language phenomena with fMRI may prove especially fruitful in the future.

Methods

Participants. Ten volunteers (nine female, ages 18–28) gave informed consent and were paid for their participation. Data from five additional subjects were excluded from the analysis, four for poor performance during the scanning session (< 85% correct overall and <75% correct on catch trials) and one for excessive head movement (>2 mm). All participants were right-handed, monolingual, native speakers of English with self-reported normal hearing, and normal or corrected-to-normal vision.

Materials. During training, participants learned a 16-word artificial lexicon (Table 2). Eight CVCVCV items referred to eight ways of modifying a set of novel shapes. Four of these modifications caused a shape to undergo a change in motion (horizontal oscillation, vertical oscillation, expansion/contraction, or rotation), and four caused changes to the shape's surface appearance (blackening, whitening, marbleizing, speckling). These items were arranged into four pairs of cohort competitors differing only at the last vowel. For two of the cohort pairs, one member of each pair referred to a motion change (e.g., *biduko*, horizontal oscillation) and the other to a surface change (e.g., *biduka*, speckled texture). Both members of one of the remaining cohort pairs referred to motion changes, and both members of the final cohort pair referred to surface changes. Eight CVCV lexical items referred to eight novel shapes not nameable with a single English word. The shapes were perceptually distinct and uniformly gray in color before modification. All of the shape names were phonologically unique.

The words were spoken as isolated tokens by a male native speaker of English and were recorded in a quiet room on a Kay Elemetrics Computerized Speech Lab (model 4300B) sampling at 22,050 kHz, for use with ExBuilder, a custom software package for stimulus presentation. The mean duration of the trisyllabic items was 902 ms (range 891–907 ms), and the mean duration of the bisyllabic items was 702 ms (range 699–706 ms). Average sound duration was 1,604 ms for the combined change and object descriptions (range 1,590–1,613 ms). There was no systematic length bias along any competitor dimension.

Training. Participants were trained on the pairings of lexical items with objects and events for 1 hour per day on each of the 3 days before the scanning session. An observation task and a decision task were used during training. In the observation task, participants initiated the trial by clicking a central fixation cross. They then observed a scene lasting 3 s. Depending on the stage of training, the scene consisted of an object that did not change, an untrained shape (circle) undergoing modification, or an object undergoing modification. Beginning 300 ms after the onset of the visual stimulus, participants heard the description of the scene, consisting of the names of the change and object present on the screen. At the end of the 3-s interval, the screen went

blank. After 500 ms, the fixation cross reappeared and the participant could start the next trial at any time.

In the decision task, participants initiated the trial by clicking on a central fixation cross. After 300 ms, but before any other visual stimulus was present on the screen, the participant heard a two-word description of a scene, with the change name presented first. Five-hundred milliseconds after the end of the sound stimulus, an object undergoing a change appeared on the screen. The participant indicated whether the previous description matched the scene by clicking the left mouse button for "match" and the right button for "no match." Immediately after the participant's response, the correct object/change combination and the correct description were presented as feedback. The correct object/change combination remained onscreen for 3 s. There was a 500-ms delay between the end of one trial and the fixation cross appearing to signal that the participant could begin a new trial when ready.

On the first 2 days of training, participants completed multiple observation and decision task blocks. On the first day, some blocks contained only objects and some blocks contained only changes, but by the beginning of the second day, both objects and changes were present in all training trials. To familiarize participants with the testing environment, they completed four modified decision blocks on day 3 while inside a mock scanner. Throughout training, 50% of all decision trials were mismatch trials, where either the object or the change that was seen did not match the spoken description. Because the scanner is noisy and the pneumatic headphones provide slightly degraded speech input, participants were gradually familiarized with the testing conditions. On day 1, participants heard clear speech. On day 2, all lexical items were low-pass filtered to mimic the output of the pneumatic headphones. Day 3 training occurred in the mock magnet, where participants heard the filtered words in a background of recorded echo planar imaging (EPI) noise.

Testing and fMRI Data Acquisition. Participants completed five 32-trial runs (9 min each) of a modified decision task on the fourth day while being scanned. Trials in the modified decision runs were completely computer-paced. In addition, the percentage of no-match trials was raised from 50% to 60% of all trials so that there were enough catch trials to determine whether participants were mishearing the stimuli. A central fixation cross was onscreen whenever objects were not present. Eight seconds of data were collected before the first trial in each run to allow magnetization to reach steady-state levels. On each trial, participants first heard a scene description consisting of a two-word phrase containing the name of the change and the name of the object. At a varying interval (3, 4, 5, 6, or 7 s) after the onset of the sound, but in all cases after the two-word phrase had finished, the fixation cross disappeared and the visual scene, consisting of an object undergoing a change, appeared. This scene persisted for a varying interval (2, 2.5, 3, 3.5, or 4 s) before it was replaced by the fixation cross. When the cross reappeared, participants indicated their response by using a button box held in the right hand. Participants used the same two fingers they had used to make mouse clicks during training. Participants then received auditory feedback in the form of a "ding" sound if correct or a "buzz" sound if incorrect. Participants also were buzzed if they did not respond within 3 s and the trial was counted as incorrect. After the 3-s response window had ended, there was a varying (1, 3, 5, 7, or 9 s) ITI between trials.

After the experimental runs, an MT localizer scan was performed. While participants fixated a central cross, alternating 20-s blocks of moving and stationary dots were presented for a total of 200 s. During the motion blocks, dots moved radially at 7°/s in an annulus ranging from 1° to 14°.

Whole-brain imaging data were collected by using a Siemens Trio 3T scanner and eight-channel head coil. EPI scan parameters were as follows: repetition time (TR), 2 s; echo time (TE), 30 ms; flip angle, 90°; with 30 axial slices (4 mm thick, 0 mm skip) in a 256-mm² field of view, with 4-mm³ isotropic voxel size. In addition, a 3D anatomical image was collected by using a T1-weighted MPRAGE sequence (TR = 2,020 ms, TE = 3.93 ms, inversion time TI = 1,100 ms, 1 mm³ isotropic voxel size, 256 × 256 matrix) reconstructed into 160 sagittal slices. Visual stimuli were displayed by using a JVC DLA-SXS21E projector and were presented on a rear-projection screen placed at the end of the bore of the magnet (viewing distance = 0.8 m) and viewed with a mirror mounted above the eyes at an angle of ≈45°. Auditory stimuli were presented through Etymotic ER-30 insert headphones, with additional hearing protection provided by an earmuff modified to permit the pneumatic tubes and foam earpieces to pass through. Head motion was minimized with the use of foam padding around the head and neck.

fMRI Data Analysis. Imaging data were analyzed by using AFNI (41). Functional images were corrected for head motion and aligned to the last time point of the last task run. Images were smoothed with an 8-mm FWHM Gaussian kernel and scaled to percent change values by dividing the signal in each voxel at each

time point by the mean signal for that voxel across the scan. These preprocessed data sets were then subjected to regression analysis. Reference time series for eight event types (auditory: motion target/motion cohort, motion target/nonmotion cohort, nonmotion target/motion cohort, and nonmotion target/motion cohort; visual: motion, nonmotion, response, and auditory error trials) were created by convolving the stimulus presentation times with AFNI's standard gamma-variate hemodynamic response function. Note that the stimulus presentation time lasted <1 TR for the auditory interval, but varied from 1 to 2 TR for the visual interval. Stimulus duration was modeled accordingly. The output parameters of the motion correction were included as regressors of no interest. The anatomical scans and the fit coefficients for each subject, converted to percent signal change, were normalized to MNI space. Separate one-way mixed-effects ANOVAs were performed on the normalized coefficients obtained from the individual subject regression for the auditory and visual intervals. Conjunction analyses were used to identify clusters significantly active for both the visual interval analysis and the moving/static dots analysis, as well as between the visual and auditory interval analyses. For

consistency, all group statistical maps were evaluated at a threshold of $P = 0.01$, uncorrected. This threshold, which would be lenient in a whole-brain analysis, was chosen based on the focus of the analyses on two relatively small and well delineated ROIs. The analyses support this approach, confirming that the foci of activation are centered at locations predicted by the existing literature for both MT/V5 and the auditory-motion region anterior to MT/V5. We recognize, however, that the robustness of the activation in MT/V5 induced by visual motion resulted in the identification of an area centered on MT/V5 but extending beyond its standard boundaries (31–33).

ACKNOWLEDGMENTS. We thank Joseph Devlin, David Kemmerer, Sheila Blumstein, Florian Jaeger, and Robert Jacobs for helpful comments and Pat Weber, Remya Nair, Rebecca Achtman, Aaron Newman, Keith Schneider, and Jennifer Vannest for help with data collection and analysis. This work was supported by National Institute on Deafness and Other Communication Disorders Grant DC-005071, Packard Foundation Grant 2001-17783, and a National Science Foundation Graduate Research Fellowship (to K.P.R.).

- Allopenna PD, Magnuson JS, Tanenhaus MK (1998) Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *J Mem Lang* 38:419–439.
- Magnuson JS, Tanenhaus MK, Aslin RN, Dahan D (2003) The time course of spoken word learning and recognition: Studies with artificial lexicons. *J Exp Psychol Gen* 132:202–227.
- Marslen-Wilson WD, Zwitserlood P (1989) Accessing spoken words: The importance of word onsets. *J Exp Psychol Hum Percept Perform* 15:576–585.
- Prabhakaran R, Blumstein SE, Myers EB, Hutchinson E, Britton B (2006) An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia* 44:2209–2221.
- Dahan D, Tanenhaus MK (2005) Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychon Bull Rev* 12:453–459.
- Creel SC, Aslin RN, Tanenhaus MK (2008) Heeding the voice of experience: The role of talker variation in lexical access. *Cognition* 106:633–664.
- Hauk O, Johnsrude I, Pulvermuller F (2004) Somatotopic representation of action words in human motor and premotor cortex. *Neuron* 41:301–307.
- Yee E, Blumstein SE, Sedivy JC (2008) Lexical-semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *J Cognit Neurosci* 20:592–612.
- MacDonald MC, Pearlmutter NJ, Seidenberg MS (1994) The lexical nature of syntactic ambiguity resolution. *Psychol Rev* 101:653–675.
- McGurk H, MacDonald J (1976) Hearing lips and seeing voices. *Nature* 264:746–748.
- Richardson DC, Spivey MJ, Barsalou LW, McRae K (2003) Spatial representations activated during real-time comprehension of verbs. *Cognit Sci* 27:767–780.
- Zwaan RA, Madden CJ, Yaxley RH, Aveyard ME (2004) Moving words: Dynamic representations in language comprehension. *Cognit Sci* 28:611–619.
- Meteyard L, Bahrami B, Vigliocco G (2007) Motion detection and motion verbs: Language affects low-level visual perception. *Psychol Sci* 18:1007–1013.
- Chao LL, Haxby JV, Martin A (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci* 2:913–919.
- Kable JW, Lease-Spellmeyer J, Chatterjee A (2002) Neural substrates of action event knowledge. *J Cognit Neurosci* 14:795–805.
- Kable JW, Kan IP, Wilson A, Thompson-Schill SL, Chatterjee A (2005) Conceptual representations of action in the lateral temporal cortex. *J Cognit Neurosci* 17:1855–1870.
- Martin A, Haxby JV, Lalonde FM, Wiggs CL, Ungerleider LG (1995) Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 270:102–105.
- Martin A (2007) The representation of object concepts in the brain. *Annu Rev Psychol* 58:25–45.
- Wallentin M, Lund TE, Ostergaard S, Ostergaard L, Roepstorff A (2005) Motion verb sentences activate left posterior middle temporal cortex despite static context. *NeuroReport* 16:649–652.
- Noppeny U, Friston KJ, Price CJ (2003) Effects of visual deprivation on the organization of the semantic system. *Brain* 126:1620–1627.
- Noppeny U, Josephs O, Kiebel S, Friston KJ, Price CJ (2005) Action selectivity in parietal and temporal cortex. *Cognit Brain Res* 25:641–649.
- Vigliocco G, et al. (2006) The role of semantics and grammatical class in the neural representation of words. *Cereb Cortex* 16:1790–1796.
- Kan IP, Barsalou LW, Solomon KO, Minor JK, Thompson-Schill SL (2003) Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognit Neuropsychol* 20:525–540.
- Weisberg J, van Turrenout M, Martin A (2007) A neural system for learning about object function. *Cereb Cortex* 17:513–521.
- James TW, Gauthier I (2003) Auditory and action semantic features activate sensory-specific perceptual brain regions. *Curr Biol* 13:1792–1796.
- James TW, Gauthier I (2004) Brain areas engaged during visual judgments by involuntary access to novel semantic information. *Vision Res* 44:429–439.
- Gennari SP, MacDonald MC, Postle BR, Seidenberg MS (2007) Context-dependent interpretation of words: Evidence for interactive neural processes. *NeuroImage* 35:1278–1286.
- Kourtzi Z, Kanwisher N (2001) Activation in human MT/MST by static images with implied motion. *J Cognit Neurosci* 12:48–55.
- Goebel R, Khorram-Sefat D, Muckli L, Hacker H, Singer W (1998) The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *Eur J Neurosci* 10:1563–1573.
- Tootell RB, et al. (1995) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15:3215–3230.
- Dumoulin SO, et al. (2000) A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cereb Cortex* 10:454–463.
- Dukelow SP, et al. (2001) Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J Neurophysiol* 86:1991–2000.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22:7195–7205.
- Damasio H, et al. (2001) Neural correlates of naming actions and of naming spatial relations. *NeuroImage* 13:1053–1064.
- Tranel D, Martin C, Damasio H, Grabowski TJ, Hichwa R (2005) Effects of noun-verb homonymy on the neural correlates of naming concrete entities and actions. *Brain Lang* 92:288–299.
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC (1995) Integration of visual and linguistic information in spoken language comprehension. *Science* 268:1632–1634.
- Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN (2007) The dynamics of lexical competition during spoken word recognition. *Cognit Sci* 31:1–24.
- McMurray B, Tanenhaus MK, Aslin RN (2002) Gradient effects of within-category phonetic variation on lexical access. *Cognition* 86:B33–B42.
- Revill KP, Tanenhaus MK, Aslin RN (2008) Context and spoken word recognition in a novel lexicon. *J Exp Psychol Learn*, in press.
- O'Craven KM, Kanwisher N (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cognit Neurosci* 12:1013–1023.
- Cox RW, Hyde JS (1997) Software tools for analysis and visualization of fMRI data. *NMR Biomed* 10:171–178.