Figure 1. Experimental design and visual stimuli. *A*, Stimuli used in the main experiment. Transparent condition: the target rotated behind two transparent occluders. fMRI activity for this stimulus is used as a baseline to assess amodal completion-related activity. Divided condition: the divided target alone rotated around the central fixation. This stimulus simulates fragmentation of visual elements due to occlusion in natural scenes. Occluded condition: the target rotated around the central fixation passing behind the occluders. In this configuration, both spatial image features, such as T-shaped junctions and temporally preceding experience with seeing, the complete appearance of the target before partial occlusion promote amodal perception. Nonoccluded condition: the divided target rotated so as to overlap two stable occluders. In this configuration, when the divided target rotated and overlapped with one of the occluders, T-shaped junctions over the occluder promote occlusion perception, although the observer who sees whole appearance of the target before overlapping knows that the target is divided, never occluded. *B*, Schematic view of target rotation (occluded condition). Participants viewed continuous rotation of the target while fixating the central dot.

(Nakayama et al., 1995). This "bottom-up" hypothesis may be supported by physiological findings that some neurons in V1 or V2 respond to an occluded bar with crossed disparity, so that it appeared to be behind the occluder (Sugita, 1999; Bakin et al., 2000). However, the early topographic hypothesis lacks direct human neural evidence. Although a pioneering human fMRI-adaptation study reported that V1 contributes to amodal completion (Rauschenberger et al., 2006), it could be argued that the observed V1 adaptation may not reveal the reconstructed representation of the occluded portion, and instead may simply reflect adaptation to low-level image features, such as T-shaped junctions at occluding borders. To resolve these disparate accounts is important to understand how the visual system reconstructs coherent percepts from fragmentary and incomplete retinal inputs.

Here, we investigated the topographic representation hypothesis in human early visual cortex. By measuring fMRI activity within the precisely defined cortical regions corresponding to the occluded portion of an object, we provide clear evidence that human early visual areas, V1 and V2, indeed represent the occluded portion as if it were pictorially completed. The observed completion-related responses cannot be simply explained by attentional boosts of fMRI responses. Furthermore, we find that the early topographic representation of occlusion is modulated by prior knowledge of the appearance of the complete object seen before partial occlusion.

Materials and Methods

Participants
Eight healthy adults (two females, one left-handed, 31-years-old on average) with normal or corrected-to-normal vision participated in the

main experiments. Eight adults participated in an additional attention control experiment (all males, one-left-handed, 28 years old on average). Three males participated in both experiments. All participants provided oral and written informed consent. All participants had experience in psychophysical experiments and were well trained to sustain eye fixation. The Ethics Committees of the Graduate School of Medicine and that of Human and Environmental Studies, Kyoto University, approved the experimental procedures in advance.

Visual stimuli and experimental design
Visual stimuli were generated on a PC using a publicly available OpenGL-compatible Visualization Toolkit (VTK) C++ library (Kitware). The stimuli were presented via a gamma-corrected (Ban and Yamamoto, 2013) digital-light-processing type projector (U2-1130, PLUS Vision) and back-projected onto a translucent screen. Participants viewed the screen through a first surface mirror tilted 45° and positioned in front of the participant's forehead.

The stimuli consisted of a continuously rotating volumetric triangular wedge (target) and two stable occluders (Fig. 1). The target wedge spanned 20° in polar angle and extended 12.5° of visual eccentricity from a point of central fixation. On the target, a continuous brick-wall texture was mapped to enhance pictorial depth perception and local luminance changes accompanying the target rotation. The occluders were uniform gray and covered 4.2–8.3° in eccentricity in the upper-right and lower-left visual field quadrants and were stable in shape and size throughout a scanning run.

In the main experiments, we manipulated the spatial and temporal configurations of the stimuli using the following four experimental conditions: transparent, divided, occluded, and nonoccluded (Fig. 1*A*). In the transparent condition, the occluders were transparently rendered using OpenGL opacity function (set to 0.5) and the target rotation behind the transparent occluders was visible. In the divided condition, the occluders were removed from display and the target was divided by removing the middle segment of the target that was to be occluded; observers viewed the divided target rotation alone. Here, this stimulus configuration may be, at a glance, perceived as the whole target rotated behind a gray (same color with the background) annulus occluder. However, the target stimulus is actually a volumetric triangular (3D) wedge projected on a 2D screen. Therefore, when it is divided and rotated along its own axis, visible triangular edges at the middle segment prevented the target from being perceived as if it is occluded by the invisible apparent occluder. All the participants actually reported that the target was divided without any occlusion percept when the stimulus was presented outside the scanner in advance of the main experiment. In the occluded condition, the target rotated behind two stable opaque occluders and the middle segment of the target was invisible when the target came to the upper-right or lower-left visual quadrant. In the nonoccluded condition, the divided target (the part that would be masked by the occluder has been removed) rotated and two stable opaque occluders were located in the upper-right and lower-left visual field quadrants. In this condition, although the target was divided, the stimulus spatial configuration was the same with occluded condition when the target passed through one of the occluders, whereas an observer seeing the sequence of target rotation knew that the target was divided. This configuration was used to test whether preceding cognitive context (observer's knowledge about the appearance of the target acquired

before the target came to overlap with the occluders) could be reflected in occlusion completion-related fMRI responses.

In all stimulus conditions, the target began to emerge from slightly below the right horizontal visual meridian and rotated smoothly around a central fixation point in a counterclockwise fashion at a rate of 10 deg/s (Fig. 1B). To enhance the impression of occlusion by pictorial and motion cues, we also put the target rotation around its own axis at 120 deg/s (see arrows in Fig. 1B). The target rotation around the central fixation (36 s/rotation) was continuously repeated 6 times in an fMRI scanning run. An fMRI scanning run thus lasted 216 s, plus a 10 s adaptation period at the beginning of each scan to prevent startup magnetization transients.

Participants viewed the target rotations sustaining eye fixation on the central fixation point. Control of participants' attention and blocking of effects of eye movement were also of crucial importance for the present study. We therefore imposed an attention-demanding task related to the central fixation point. Specifically, the color of the fixation point changed randomly (at 1 s intervals, 500 ms duration) to red, green, blue, or yellow, and then turned back to pale blue (default color). We instructed participants to press an optical switch when the color of the fixation point turned red. Image data with <80% correct responses to the fixation task were discarded.

To discriminate effects of attentional boosts on fMRI activity from the amodal completion-related responses, we further conducted an additional control experiment with manipulating the central fixation task difficulty. In this control experiment, the color of the central fixation changed more frequently (at every 500 ms, 200 ms duration) and participants were asked to concentrate on the same color detection task through all scans. The ratio of correct responses for this experiment was 60–80% and all the data were used.

These and the other stimulus parameters such as size, luminance, and speed of the target rotation were set to optimize the perceptual salience of amodal completion on participants' reports outside the scanner, and for clear and reliable determination of cortical retinotopic representation. Further, the luminance contrast of the occluders from gray background was set relatively high (120 cd/m^2 for the uniform gray background, 194 cd/m^2 for the occluders) as to avoid the Troxler effect (Friedman et al., 1999; Von der Heydt et al., 2003; when steady fixation is maintained, the contrast of an object in the peripheral visual field gradually decreases, and the object finally becomes invisible). The continuous target rotation sweeping the entire visual field repeatedly also minimizes such effects.

The "amodal" nature of occlusion perception, which is accompanied by no concurrent sensory representation such as luminance enhancement of the occluded portion, makes it difficult to measure perceptual performance quantitatively. However, through questionnaire and actual stimulus demonstrations on computer displays outside the scanner in advance of the scanning session, all participants reported occlusion percepts for occluded and not for divided condition. In nonoccluded condition, when the target overlapped with the occluders, all participants reported occlusion-like percepts promoted by spatial image features, such as T-shaped junctions at the occluding borders, but nonetheless reported that they knew the target was divided, and not occluded.

Imaging data acquisitions
Participants were scanned on 3.0 tesla Siemens MAGNETOM Trio Scanner equipped with an eight-channel phase-array head coil (Siemens) at the Graduate School of Medicine, Kyoto University, Japan. Blood oxygenation level-dependent (BOLD) contrast was obtained with a gradient-echo echo-planner imaging (EPI) sequence (TR 2000 ms, TE 30 ms, flip angle 90°, matrix size 64 × 64, voxel size 3 × 3 × 3 mm^3 without gap). The scanned volume included 24 slices nearly perpendicular to the calcarine sulcus. The onset of stimulus presentation and fMRI scanning were synchronized using a custom-made triggering system. Data acquisitions were repeated at least four times for each stimulus condition in a day. T1-weighted 3D high-resolution (voxel size 0.94 × 0.94 × 1.0 mm^3, 208 axial slices) anatomical images were also acquired for each participant to allow accurate spatial coregistration between functional and anatomical image data, cortical gray/white matter segmentation using mrGray software (Teo et al., 1997), and cortical surface reconstruction (Yamamoto et al., 2008, 2012).

Reconstruction of the cortical surface
The cortical surfaces of the individual participants were reconstructed from the structural T1-weighted image volumes. The detailed procedures are described previously (Yamamoto et al., 2008, 2012). Briefly, we generated the surface lying approximately in the middle of the gray matter using a method that was a hybrid of volume segmentation (Drury et al., 1999) and surface deformation (Dale and Sereno, 1993) techniques. First, the voxels that belonged to the cortical gray matter were segmented from the rest of the volume using mrGray software. mrGray enabled us to identify the white matter, CSF, and three layers of gray matter semiautomatically. The segmented gray matter was ~3 mm thick and they were classified the first, second, and third layers from the gray-white matter boundary. The segmented volume was then smoothed using a 3D Gaussian filter to minimize any nonbiological irregularities. Next, a surface representation was created for the gray-white matter boundary. At this point, we did not create the surface for the middle of the gray matter to minimize topological defects, such as bridges between cortical sulci. We then computed a concrete mean voxel value for the first gray matter layer in the smoothed segmented volume and applied the marching cube algorithm (Lorensen and Cline, 1987), which extracted an isosurface tessellated with ~300,000 triangles. The number of triangles was then reduced to 200,000 using the decimation algorithm (Shroeder et al., 1992). Finally, the triangulated surface was deformed such that it lay in the middle of the gray matter by relaxing it against the smoothed segmented volume. We used the deformable template algorithm (Dale and Sereno, 1993) for deformation. Finally, the resultant surface was visually inspected for positional accuracy and topological errors by overlapping it with the original structural MR volume. An additional method used to detect topological defects was extensive smoothing of the surface, which highlighted the surface defects as sharp edges. If the surface was inaccurate or had severe defects, corrections were made in mrGray manually and the subsequent procedures were repeated. We also created inflated, that is, hyper-smoothed versions of the reconstructed surfaces. We used an inflation algorithm that was essentially similar to that proposed by Fischl et al. (1999). Each participant's functional data were coregistered into individual reconstructed surface space. Reconstructed cortical surfaces were used to visualize cortical activity, to identify retinotopic visual areas, and to sample functional voxels in the target regions of interest.

fMRI data analyses.
Preprocessing. For all fMRI data, the first five volumes were discarded to prevent startup magnetization transients. All data underwent slice-timing correction and 3D motion correction (Woods et al., 1998). Slow drifts in signal intensity were removed by a classical decomposition technique as all data contained periodic responses. Then, the imaging data were coregistered to the anatomical volume. All data were analyzed in these individual anatomy spaces. No further spatial/temporal smoothing and spatial normalization (e.g., Talairach transformation) were applied.

Identifying retinotopic visual areas. Before conducting the main experiments, we performed retinotopy localizer scans (at least four scans for each of polar and eccentricity maps) and identified 12 visual areas, V1d, V1v, V2d, V2v, V3d, V3v, V3A, V3B, V4v, V7, LO (lateral occipital region), and MT+, following the standard phase-encoded analysis technique (Engel et al., 1994; Sereno et al., 1995; DeYoe et al., 1996) for each participant (Fig. 2). We used luminance and chromaticity flickering (10 Hz) checkerboard patterns for these retinotopy scans. The details were described previously (Yamamoto et al., 2008, 2012). The polar angle map allowed us to identify the borders of V1d(v)/V2d(v), V2d(v)/V3d(v), V3d/V3A (Tootell et al., 1997), and V3A/V7 as reversals in the polar angle and field-sign map. Eccentricity map was also used to identify V3A, V3B, and V7 borders as the foveal representations of these areas were displaced superiorly with the confluent foveal representation of areas V1, V2, and V3 (for a review, see Wandell et al., 2005, 2007). We identified V3B (Smith et al., 1998) so that it was located just anterior to V3d and its foveal representation was just adjacent to the V3A foveal representation. We identified MT+ as the region within the dorsal posterior limb of the inferior temporal sulcus, which had a crude eccentricity map with predominantly foveal representation superiorly and peripheral representation inferiorly (Huk et al., 2002). We confirmed that this region mostly
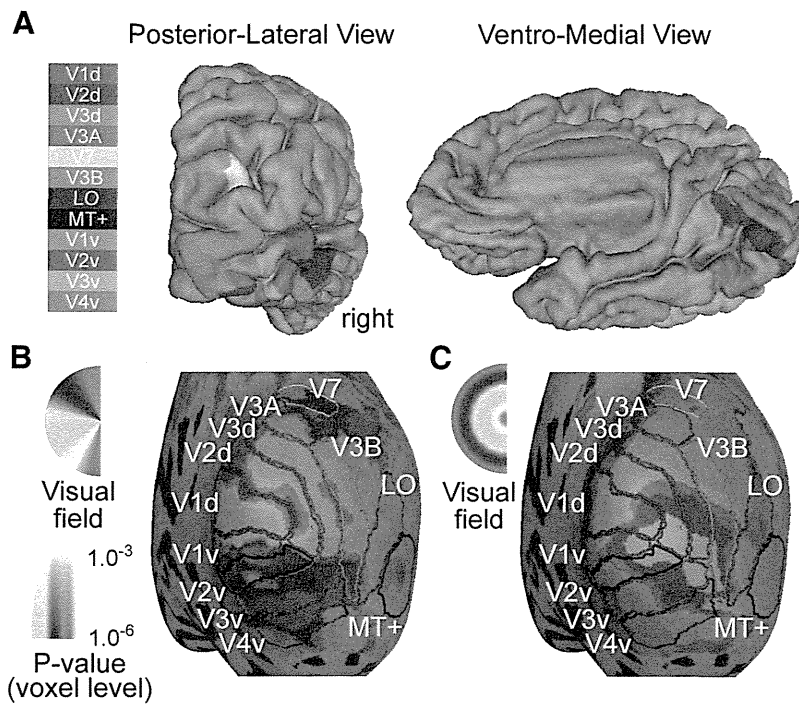
Figure 2. Identified retinotopic visual areas. Twelve retinotopic regions of interest. *A*, Locations of the visual areas V1, V2, V3, V4v, V3A, V3B, V7, LO, and MT+ in one subject's right hemisphere from posterior lateral view (left) and ventromedial view (right). The icon to the left of the panel indicates the relationship between color and visual area. Borders of the areas were determined from the polar-angular (*B*) and eccentricity (*C*) visual field representations measured by separated phase-encoded retinotopic mapping experiments. The color overlay on the inflated cortex indicates the preferred stimulus angle or eccentricity at each cortical point, and the colored lines indicate each area's border. The icons to the left of each panel indicate the relationship between color and visual field position.

Phase-encoded analysis to evaluate cortical retinotopic activity corresponding to the occluded portion. Target rotation around the central fixation point caused periodic shifts of neural activity in the early visual areas; because their neurons have limited receptive fields whose centers are organized to form a continuous mapping between the cortical surface and the visual field, a portion of the early visual cortex responded periodically when the target swept across the corresponding visual field location (Wandell et al., 2007). We investigated whether we could observe this periodical shift of activity even when the target stimulus rotation was partly occluded. To evaluate retinotopic periodic responses, the data were analyzed using a Fourier transformation-based method, well documented in the standard retinotopy studies (Engel et al., 1994; Sereno et al., 1995; DeYoe et al., 1996). First, a Discrete Fourier transformation was computed for the time series of each voxel after converting the raw signals to percentage signal changes. Then, statistical testing to estimate the significance of correlation of BOLD signal at the stimulus frequency (1/36 Hz) was performed by comparing the squared amplitude at the stimulus frequency with the sum of squared amplitudes at the other frequencies, which yielded an F-ratio. The F-ratio was converted to a P value considering degrees of freedom of the signals (number of time points). In mapping retinotopic responses on the cortical surface, the polar angle phase of each significantly activated voxel was displayed using a continuous color scale. Here, a voxel-level Fourier F-statistic value, $F_{(2,105)} > 7.38$ and $p < 0.001$ (uncorrected), was used as a criteria of significance for all Fourier-based analysis (e.g., mapping activity on the cortical surface) if there is no further information.

Mapping fMRI responses on reconstructed cortical surface. We visualized imaging data by on the reconstructed cortical surfaces generated following the procedures described above. The surface representation was created by sampling the voxel responses at each node of the surface mesh within a radius of 4 mm and averaging them. The representation was color-coded by assigning different colors to different response phases, and by altering the saturation of colors such that each color increased saturation corresponded to a statistical significant level.

Subdivision of ROIs and detailed voxel-based analysis. To evaluate the periodic fMRI responses quantitatively and to investigate the spatial specificity of the periodic responses, we further performed detailed voxel-based analyses. To this end, each of the retinotopic ROIs (subregions in V1–V3 that corresponded to the occluder retinotopic positions) were further divided into three subregions along cortical visual-field polar angle or eccentricity representation for each participant. fMRI responses were separately resampled within these subregions and compared. Visual-field polar angle representations were determined based on cortical activity evoked by the transparent stimulus (Fig. 1A). Eccentricity representations were determined by an expanding ring stimulus used in the standard retinotopic experiments. fMRI voxel responses were resampled at each node of these sub-ROIs' meshes within a radius of 2.5 mm. Note that, in all the voxel-based detailed analyses, all voxels located along the border of each ROI (~2 mm radius from the borders) were carefully removed in advance because it is possible that these voxels contain BOLD responses from neighboring regions where target rotation is completely visible. We thus compared neural activity only within the strictly limited regions of predefined ROIs that retinotopically represent regions covered by the occluders. The number of voxels survived in this preprocessing were 21.94 ± 9.03 (mean and SD) in V1 sub-ROI, 15.06 ±

overlapped with the middle temporal region, and exhibited a strong response to motion stimuli. We refer to the large fan-shaped region between areas V3B and MT+ as LO (Malach et al., 1995), which had a relatively clear polar and eccentricity representations (Levy et al., 2001; Larsson and Heeger, 2006). There is an enduring dispute regarding subdivision of the ventral occipital cortex anterior to V3v (Wandell et al., 2005, 2007). Here, we identified an area, V4v, after Hadjikhani et al. (1998) stressing consistency not in the angle but in the eccentricity map in our data (Yamamoto et al., 2012). V4v in this study roughly corresponded to the posterior half of hV4 (Zeki et al., 1991; Bartels and Zeki, 2000; Brewer et al., 2005). Note that we also analyzed data of some participants by reidentifying these regions as hV4 and VO1/VO2 using different criteria (Brewer et al., 2005), but we could not find any differences of the results in the present study.

Identifying regions of interest. Then, we further localized the cortical subregions in each retinotopic area (V1, V2, V3, and V3A) that corresponded to the locations of the occluders presented in the subsequent main experiments (see Fig. 4A, bottom left). The localized regions were used as regions of interest (ROIs) for the analyses of the main experiments. These ROIs were identified in separate block-design scans for each participant and each scanning day (two runs per participant and per day) by presenting flickering checkerboards in the same locations of two occluders used in the main experiment alternating with the compensating checkerboard pattern. To identify the occluder regions, we applied the standard GLM analysis (occluder position vs the compensating pattern) on the data and extracted voxels from each visual area that showed responses at $p < 0.001$ (uncorrected) for the checkerboard pattern corresponded to the occluder positions. We also localized foveal and peripheral regions that were external to the occluders for each early visual area. For V3A, we could only roughly localize two subregions, corresponding to the occluders and the external regions because V3A had a hemivisual field representation and retinotopy was not as clear as that observed in V1–V3 (Fig. 2B, C; see Fig. 4A, bottom left).

7.03 in V2, and 13.25 ± 3.45 in V3. For some sub-ROIs of a few participants, we omitted the data from the detailed voxel analyses since we could not segregate the ROIs into three subregions with precision. Thus, statistical tests in the detailed sub-ROI analyses were performed with $n = 6$ in V1, V2, and V3, whereas the other tests were performed with $n = 8$ in all ROIs.

Even after these retinotopy-based restrictions of voxels, it would be difficult to completely exclude the periodic responses from the outer regions to the occluder ROIs due to BOLD spreads. To rule out this confound more strictly, we further restricted the sub-ROIs by excluding any voxels that showed significant responses to the divided stimulus in the same scanning session. Specifically, after defining the occluder ROIs, we extracted all voxels that showed $F_{(2,105)} > 3.08$ and $p < 0.05$ (uncorrected) for the divided condition in some additional analyses.

Comparisons of voxel-based periodic activities among conditions. To compare periodic response magnitudes evoked by the visible and invisible target rotations, we calculated a relative periodicity index following the same procedures of Fourier F test described above. Specifically, we first computed voxel-by-voxel periodicity for each stimulus condition, by dividing the squared amplitude at the stimulus frequency (1/36 Hz) with the sum of squared amplitudes at the other frequencies. Here, the value indicates how strongly a voxel responds to the invisible (occluded or divided) target. Then, to directly compare periodic powers across conditions, the calculated periodicity of each voxel for the occluded condition was divided by the corresponding voxel periodicity for the divided or nonoccluded condition. These values were averaged across voxels, yielding a relative periodicity index in each ROI.

Completion index. We calculated completion index for each participant to directly compare the similarity of neural activity evoked by the invisible target rotation with that evoked by visible rotation. The completion index was defined as

$$Completion\_Index = \frac{1}{n} \times \sum_{i=1}^{n} \left(\frac{\vec{a}_i \cdot \vec{b}_i}{|\vec{b}_i|} \times \frac{1}{|\vec{b}_i|} \right),$$

where n is the number of voxels within a ROI, $\vec{a}$ is a vector consisting of the response amplitude and phase angle of a voxel for the occluded (or divided) stimulus ($\vec{b}$ = amplitude, phase), and $\vec{b}$ is a vector for the corresponding voxel response to the transparent stimulus. This index is 0 when there is no response to the invisible target rotation, negative when the responses to the invisible rotation are opposite to the visible responses, positive (usually <1) when the responses for the invisible rotation is similar with the visible rotation, and equal to 1 when the invisible target evoked exactly the same response as the visible target did. In calculations, data from the two hemispheres were averaged, because no significant difference was observed between responses in the retinotopic regions corresponding to lower-left and upper-right visual field quadrants. Unlike the periodicity analysis described above, which took into account only powers of periodicity at the target frequency (1/36 Hz) and did not consider response phases, this correlation-like analysis considered both phase and periodic power components simultaneously. This indexing analysis thus involves a more conservative approach to investigating completion-related activity.

Results

Experimental design

The target visual stimulus was a volumetric triangular wedge that continuously rotated around a central fixation point (36 s per rotation, 6 rotations per scanning session) while rotating around its own axis (Fig. 1B). In the scanner, participants viewed the target rotation while maintaining fixation on the center of the screen and performing a color detection task on the fixation point. Rotation of the target around the central fixation causes periodic shifts of neural activity as to sweep the retinotopic visual cortex as shown in standard retinotopic studies (Engel et al., 1994; Sereno et al., 1995; DeYoe et al., 1996). We used this property of periodic activity to access amodal completion-related neural responses. That is, we explored how the periodic activity corresponding to the target rotation was modulated when perception of the target was changed by systematically manipulating the target appearance and/or its surrounding spatial and temporal contexts leading to modal or amodal completion. We reasoned that visual areas whose activity reflects the topographic interpolation of amodal completion, if they exist, should exhibit robust periodic responses even to the occluded rotation in the absence of direct sensory input.

fMRI responses to a visible object in retinotopic visual areas

We first measured cortical activity with completely visible target rotation as a baseline for each participant to evaluate amodal completion-related activity for subsequent stimuli. Here, to match stimulus configurations, such as edges with the stimuli described below, we additionally presented static transparent occluders in the upper-right and lower-left visual field quadrants (Fig. 1A, transparent condition). This visible target rotation caused reliable periodic shifts in neural responses in overall visual eccentricity spanning the rotating target.

Here, it is possible that the transparency of the occluders may cause systematic shifts in response phases (delays) in retinotopic areas that result in inadequate evaluation of amodal completion-related activity. To rule out the possibility in advance, we compared phase-encoded responses evoked by transparent stimulus with those evoked by target rotation alone. As shown in Figure 3, the two stimuli yielded similar response patterns in early retinotopic visual areas without any response biases due to transparency. We therefore conclude that transparency of the occluders did not affect the following analyses.

Next, we explored how the observed retinotopic activity was modulated when the target was divided into two separate portions by removing the middle segment (Fig. 1A, divided condition), simulating virtual fragmentations and separations of an object that are often occur in natural scenes due to occlusion. In all participants, division of the target caused division of activation pattern on the cortical surface in V1 and V2 (Fig. 4A, top, left and right). The missing region of fMRI activity was confined to retinotopic subregions corresponding to the divided segment, which had been separately identified by additional localizer scans for each participant (Fig. 4A, bottom left; see Materials and Methods).

The separation of cortical representations of the separated visual elements has the following important implications for object representation in early retinotopic areas. First, in these areas, an object is represented by neurons whose receptive fields are too small to capture both of the divided visual elements. Second, although the separated visual elements undergoing synchronous motion induced a percept of them as a single virtual wedge-like object, the representation corresponding to such perceptual integration and completing the missing region were not used in these areas (but see Meng et al., 2005). Third, in our experimental conditions, in which participants performed a color detection task on the central fixation point through a scanning session, the representation did not include attentional (Somers et al., 1999; Sasaki et al., 2001) or imagery (Slotnick et al., 2005) effects radially spreading along the divided segments that could result in neural completion of the missing portion. Together, these findings suggest that early visual areas generally respond to only visible portions of objects based on direct sensory inputs from retina to cortex.
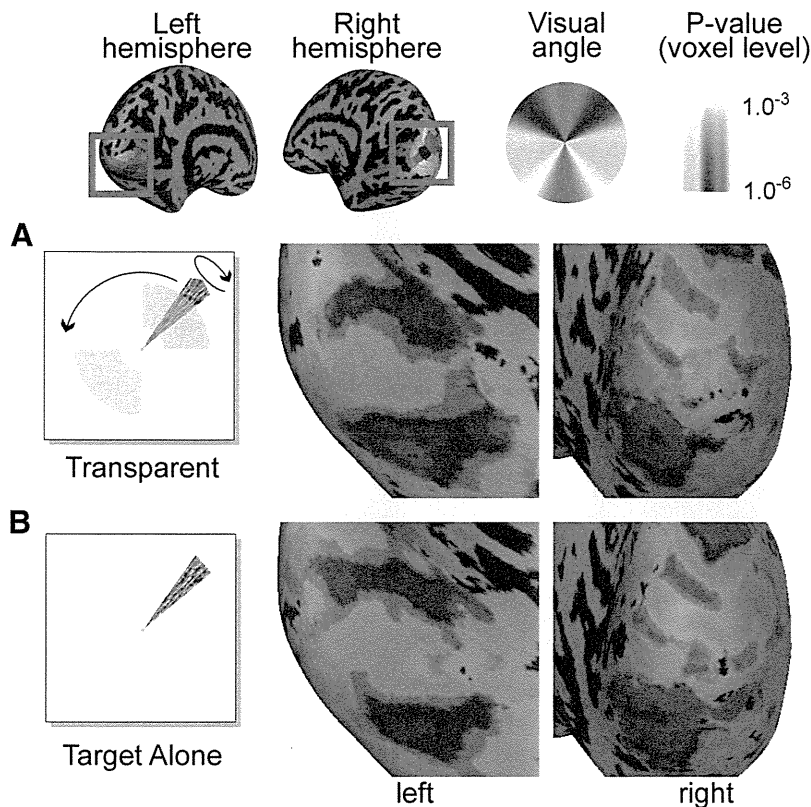
Figure 3. Representative cortical activity in response to transparent and real target rotations. Response phase angles of significantly (Fourier F test, $p < 0.001$, voxel level) activated voxels projected on participant's inflated occipital cortex. A, Phase-encoded responses for target rotation behind transparent occluders (transparent condition). B, Phase-encoded responses for target rotation alone.

Topographic fMRI responses to an occluded object

Then, what neural activity is observed in retinotopic early visual areas when the target rotation is partially occluded by an opaque occluder? We measured cortical activity while participants viewed the partially occluded target rotation (Fig. 1A, occluded condition). Note that, in this configuration, the target rotation is completely divided, as in the divided condition, and that the middle segment of the target is completely invisible while the target passes behind the occluders. Thus, if early visual areas truly represent only visible portions of an object, the pattern of cortical activation by this occluded rotation should again be divided, and no periodic neural response should be observed within the retinotopic subregions representing the occluders.

However, we found clear periodic shifts in activity corresponding to the occluded and invisible target rotation, as if the gap between the visible portions of the target had been bridged despite lack of direct sensory input to these regions. Although the responses were not as strong compared as those evoked by the transparent stimulus, V1 and V2 exhibited robust activation even for the occluded portion (Fig. 4A, bottom right; B, averaged time courses from three subportions in V1, V2, and V3). Because similar completion-like activity was not observed or weaker (Fig. 4A, top) in the divided condition, the observed periodic activity cannot be simply explained by fMRI signal leakage from external (foveal/peripheral) regions of the occluders, where rotation was completely visible. Furthermore, because the occluders were presented stable in specific locations during a scanning run and our analysis is targeting periodic activity (signal power at 1/36 Hz included in the time series), the observed periodic activity in the

occluded condition cannot be explained as an effect of the opaque occluders alone. The observed activity thus may reflect topographic activity to the invisible part of the target stimulus; geometric relations and interactions between the target and occluders may allow visible portions of an object to be connected under occlusion and may promote completion-related neural responses.

We need to note, however, that those simple analyses cannot completely rule out the alternative possibilities. We generated the cortical response maps by sampling voxels within relatively large space (4 mm radius from each cortical surface node). Therefore, there is no evidence that responses within the occluder ROIs were really free from the response confounds from visible regions. In addition, our phase-encoded experiment paradigm might work so as to bias the averaged time courses due to the effect of negative BOLD signals, etc. Specifically, it might be possible that the divided condition may show some opposite response profile (in Fourier phase domain) compared with the transparent condition. In that case, responses of the occluded condition would become relatively larger than the divided even without any actual completion-related activities. Then, we may mistakenly conclude that seemingly larger activity for the occluded condition would be related to occlusion completion. To rule out these possibilities and to confirm the topographic completion-related responses, we conducted more conservative and detailed analyses. In the following analyses, voxels located along the border of each ROI were carefully omitted. Further, the analysis was limited to voxels that were significantly activated by the ROI localizer (voxel-wise t test; $p < 0.001$, uncorrected) and that exhibited significant retinotopic responses (Fourier F test; $F_{(2,105)} > 7.38$, $p < 0.001$, voxel level) for the transparent stimulus (see Materials and Methods).

First, to test whether the occluded target rotation actually caused periodic responses to complete the invisible portion, we explored the relationship of voxel-by-voxel Fourier powers and phases between the transparent and occluded/divided conditions in V1–V3 subregions that retinotopically represented the occluder positions. If the occluded condition actually caused completion-related responses, these response phase profiles in ROIs would be similar with those of the transparent condition. The results were scatter-plotted in Figure 5. In these plots, larger dots indicate larger periodic powers (and thus smaller p value in Fourier F-statistics). By visual inspection, we found clear correlations of response phases between the occluded and transparent conditions. Furthermore, we could confirm that no voxel with opposite phase and strong periodicity was included in the occluded condition (Fig. 5A). Therefore, we can conclude that the occluded condition actually caused positive completion-like activity to the invisible target rotation. Note that we also found some voxels that had strong periodicities corresponded to the invisible target rotation even for the divided condition. However,

the total number of these periodic voxels was smaller than that of the occluded condition (Fig. 5B). The observed strong activity even in the divided condition would be probably due to inevitable BOLD spreads from outer regions. We conducted further detailed analyses to exclude these effects later.

To quantitatively evaluate these periodic responses, we directly compared powers of periodic responses (periodicity) evoked by the occluded stimulus with those by the divided, focusing on the responses within the predefined ROIs in V1–V3 representing the occluded portions (Fig. 6; see Materials and Methods). Here, higher periodicity of a voxel indicates stronger response at the stimulus frequency (1/36 Hz) corresponding to invisible target rotation. Calculated relative periodicities (periodicities of the occluded condition was divided by that of the divided) revealed that V1–V3 exhibited significantly higher completion-related periodic responses for the occluded stimuli (paired t tests to access whether the relative periodicity is >1, $t_{(7)} = 2.84$, $p = 0.014$ in V1, $t_{(7)} = 5.65$, $p = 0.00039$ in V2, $t_{(7)} = 3.63$, $p = 0.0042$ in V3. All ROIs were significant at $p < 0.05$ level after adjusting p values for multiple-comparisons using Bonferroni correction method).

It might be argued that the observed completion-like activity is not derived from the resultant neural representation of the occluded region, and instead simply reflects responses to edges generated at the boundaries of occlusion. To test this possibility, we examined the temporal and spatial specificity of the completion-like activity. We resampled cortical responses separately along visual field polar or eccentricity representations within three subregions in each ROI for each participant. If the observed responses reflected only the occluding edges, the periodic responses would not be observed in the middle position of each ROI. In contrast, if the responses truly reflected the neural representation of the occluded portion, the periodic responses separately resampled along visual polar/eccentricity representations would exhibit no bias in visual field location.

First, we plotted voxel time courses for the occluded stimulus resampled separately along polar visual angles. As shown in Figure 7A, responses were periodically shifted even after the target had completely entered the region of occlusion, indicating that the neural representation of the invisible rotation is updated from moment to moment without direct sensory
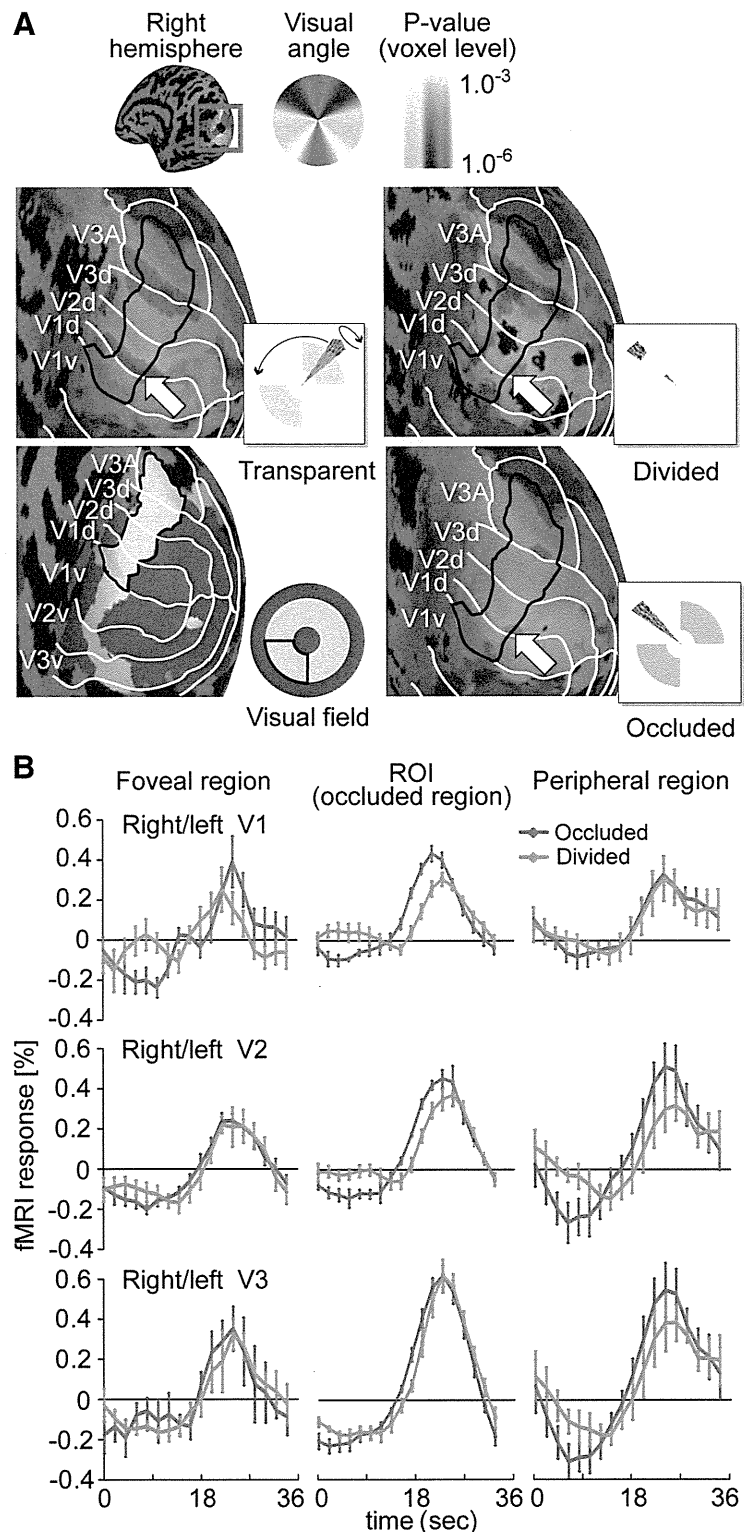

Figure 4. Representative cortical responses and fMRI time series for Occluded and Divided stimuli. **A**, Top left and right, and bottom right panels, Cortical responses projected on representative right inflated cortical surfaces (voxels with $p < 0.001$ in voxelwise Fourier F test were mapped). Colors represent the corresponding visual field locations as shown the icon above. Color saturations represent statistical p values. The black solid lines indicate retinotopic subregions representing the lower-left occluder. The white lines indicate retinotopic visual area borders. Bottom, left, Cortical activity in response to the checkerboard stimulus used for localizing the cortical subregions representing the upper-right and lower-left occluders. **B**, Averaged fMRI signal time courses evoked by the occluded and divided stimuli in V1, V2, and V3. Left/right columns show the averaged voxel time courses sampled from the foveal/peripheral regions of the retinotopic subregion corresponding to the occluded portion. Middle column show the averaged voxel time courses corresponding to the occluded portion. Here, the response phases of voxels were aligned to 16 s after the first stimulus presentation by linear interpolating and shifting the time courses voxel-by-voxel based on the response phases evoked by the transparent stimulus. Error bars, SE.
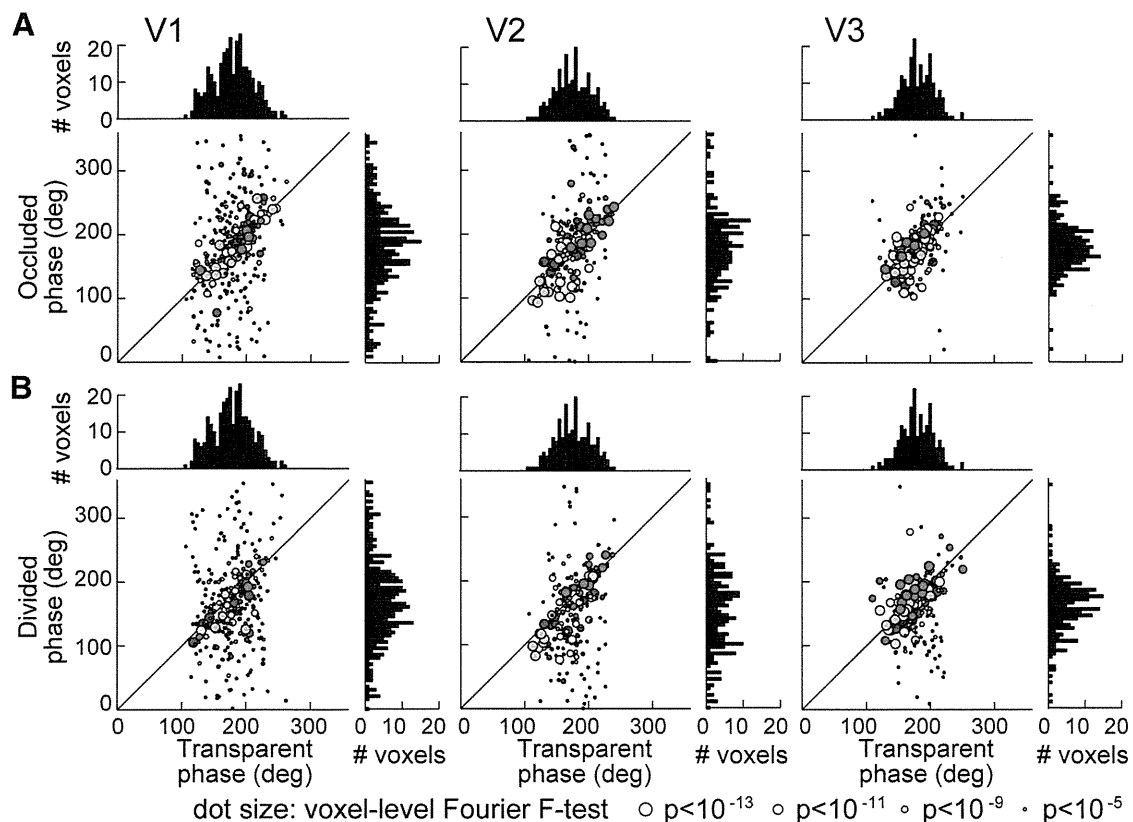
Figure 5. Relationship of response phases between the transparent and occluded/divided conditions. *A*, Voxelwise phase scatter plots, transparent versus occluded. Each dot represents individual voxel. Each color represents single participant. Dot sizes represent magnitudes of Fourier *F*-statistics (and the corresponding statistical *p* values). Larger dot indicates that the voxel contains higher-power at the target rotation frequency (1/36 Hz) compared with the sum of the powers at the other frequencies. Here, for legibility, the response phases of voxels were aligned so that the center of response phases of each ROI comes to 180° by voxel-by-voxel linear shifting based on the response phases evoked by the transparent stimulus for each participant. *B*, Voxelwise phase scatter plots, trasnparent versus divided.
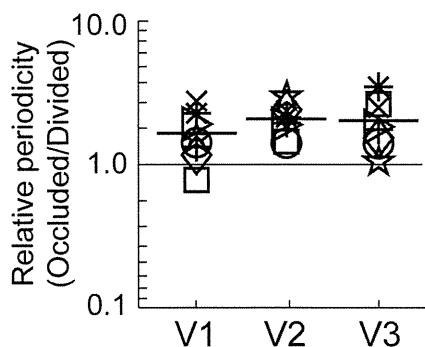

Figure 6. Relative periodicity (occluded/divided) in V1–V3 ROIs. Relative periodicity plots in V1, V2, and V3. Each shape represents a single participant. Horizontal bars represent mean values.

input. Next, we compared relative periodicities resampled along visual polar angles. As shown in Figure 7*B*, we found no bias in visual position; one-way (visual polar angles, six locations) repeated-measures ANOVA revealed no main effect of visual location ($F_{(5,30)} = 0.75, p = 0.59$ in V1; $F_{(5,35)} = 0.62, p = 0.68$ in V2; and $F_{(5,35)} = 0.43, p = 0.82$ in V3). Thus, the observed responses are not derived simply from occluding edges generated when the target enters the occluder. Further, we compared relative periodicities resampled along visual eccentricity. As shown in Figure 7*C*, the periodic responses were not limited to nearby regions of foveal/peripheral edges of the occluders (one-way

repeated-measures ANOVA revealed no main effect of visual field location; $F_{(5,30)} = 0.07, p = 0.99$ in V1; $F_{(5,35)} = 1.07, p = 0.40$ in V2; $F_{(5,35)} = 0.40, p = 0.85$ in V3). In sum, these comparisons of periodic responses over retinotopic subregions revealed that the observed activity reflects the neural completion responses.

Comparison of neural responses to amodal and real objects

We then tried to quantitatively estimate how robust the representation of an occluded object was, compared with that of its real counterpart. To this end, we calculated a completion index for each ROI (see Materials and Methods). Briefly, the index was defined as the mean of components of voxel response amplitudes evoked by an amodal surface with zero phase (delay) lags relative to the response from a real surface. If the index is 1, neural activity for the amodal rotation is completely identical to that for the visible target rotation, whereas an index of 0 indicates no completion-related response. Here, baseline neural activity of a real surface was determined separately for each participant as activity evoked by the transparent stimulus. Completion indices for the divided stimulus were also calculated for comparison. Note that this analysis can evaluate completion-related activity more conservatively than the periodicity analysis described above (Figs. 6, 7), because not only the power of periodic response but also the response phase (delay) of each voxel is taken into account. Specifically, when we only focus on periodicities of voxel responses (1/36 Hz), both positive and negative signals would equally contribute to boost periodic powers in ROIs, although
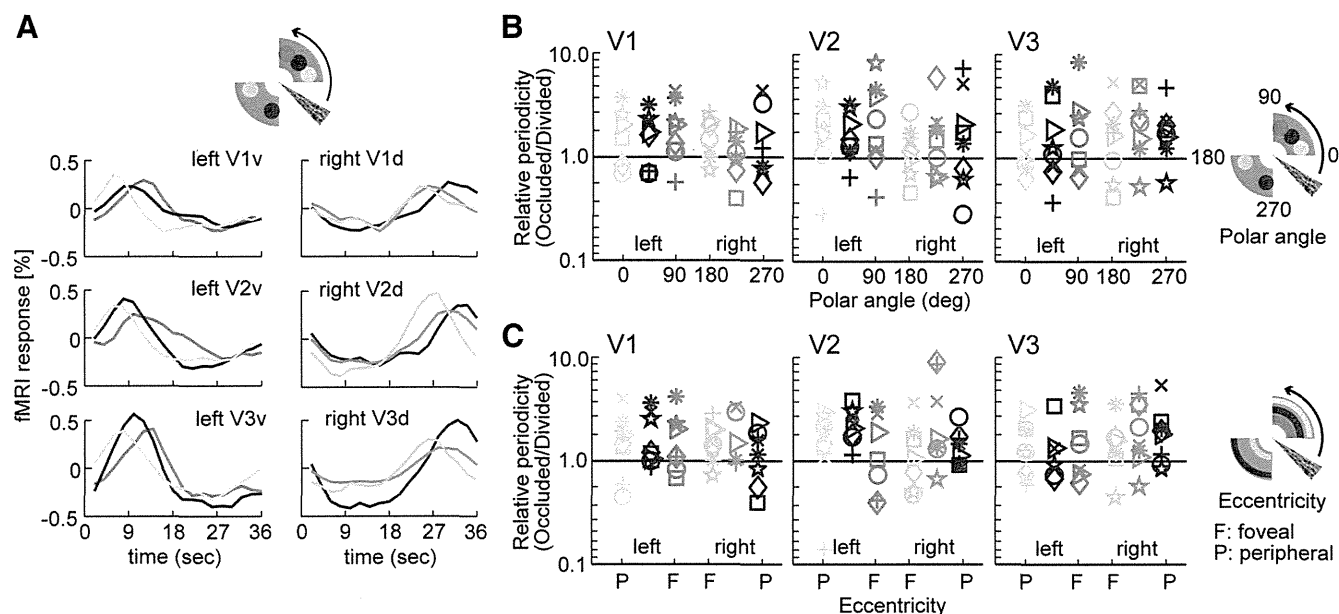
Figure 7. Temporal and spatial specificity of completion-related responses. Voxel-by-voxel responses were resampled along visual field polar angle or eccentricity representation. *A*, fMRI voxel time courses evoked by the occluded stimulus sampled and averaged separately from three subregions of each ROI along the cortical visual polar angle representations. Each color represents the corresponding visual field location. *B*, Comparison of relative periodic responses (occluded/divided) along polar angle representation. Each shape represents a single participant. Each color represents a single retinotopic position as shown in the right icon. *C*, Comparison of relative periodic responses (occluded/divided) along visual eccentricity representation.
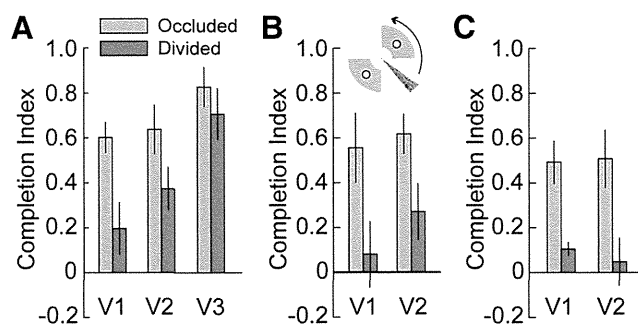

Figure 8. Completion indices in retinotopic ROIs for the occluded and divided stimuli. *A*, Completion indices in retinotopic ROIs. The indices for the occluded stimulus revealed that only V1 and V2 exhibited significant completion-related activity compared with the divided stimulus. *B*, Completion indices within the strictly limited regions in V1 and V2 corresponding to the middle position of the occluders. *C*, Completion indices in retinotopic ROIs after excluding all voxels that showed responses at $p < 0.05$ level for the divided condition to minimize BOLD spread effects in the results. The voxel exclusion was done in each scanning day and each participant separately. Error bars, SEM.

negative responses corresponding to the target rotation, if observed, would generally work to inhibit completion-related activities. By putting effect of these potential relative phase shift into the index, we can avoid this problem and evaluate amodal responses more precisely in the next analysis.

Even after the voxel response phases were taken into account, V1 and V2 ROIs (but not V3 in this analysis (paired t test; $t_{(7)} = 1.25$, $p = 0.13$) exhibited significantly higher responses to the occluded than to the divided stimulus (paired t test; $t_{(7)} = 2.75$, $p = 0.014$ in V1; $t_{(7)} = 2.96$, $p = 0.011$ in V2; significant at $p < 0.05$ level after Bonferroni correction; Fig. 8*A*). The result indicates that V1 and V2 contribute to topographic amodal completion by boosting activities for the occluded and invisible visual object. This finding also indicates that neural responses to the amodal and real surfaces have relatively similar profiles. The indices for the occluded stimulus were 0.60 in V1 and 0.64 in V2. In

contrast, although the indices for the divided stimulus were above zero level (paired t test; $t_{(7)} = 8.67$, $p = 0.000027$ in V1; $t_{(7)} = 5.77$, $p = 0.00034$ in V2; $t_{(7)} = 9.35$, $p = 0.000016$ in V3, uncorrected), their magnitudes (0.19 in V1 and 0.37 in V2) were significantly lower compared with the occluded stimulus. The amodal completion-related responses in V1 and V2 were robust even when the ROIs were strictly confined within the center positions of the occluders as a further defense against BOLD spread (paired t test, $t_{(7)} = 3.72$, $p = 0.0037$ in V1; $t_{(7)} = 3.50$, $p = 0.005$ in V2; Fig. 8*B*). Furthermore, to omit effects of BOLD spreads from the foveal and peripheral visible parts of the target stimulus more strictly, we excluded any voxels that showed significant responses (voxelwise Fourier F test, $p < 0.05$ uncorrected) for the divided stimulus from the analysis. Even after this strict limitation of voxels, V1/V2 completion-related activities were clear as shown in Figure 8*C* (paired t test; $t_{(7)} = 4.28$, $p = 0.0036$ in V1; $t_{(7)} = 3.8567$, $p = 0.0063$ in V2; significant at $p < 0.05$ level after Bonferroni correction), indicating robustness of the completion-related activity within exactly the retinotopic positions of the visual occlusion.

In foveal/peripheral regions of ROIs, no significant difference in the index was observed between conditions (paired t test; $p > 0.10$ for all the ROIs; Fig. 9*A*). This result also suggests that the effects of BOLD spreads from the outer regions into centers of ROIs would be minimal since as otherwise some differences would be also observed in the outer regions; if responses for the occluded stimulus at the occluding borders are high enough to modulate the activities in the ROIs, it would also affect the response profiles in the outer regions. Therefore, we can conclude that the neural modulation by an amodal surface is spatially limited to regions corresponding to the missing portion of an occluded object. Our previous fMRI study (Ban et al., 2006) found enhancement of retinotopic neural responses to a visible portion of a target stimulus when another stimulus was simultaneously presented to form a global structure even if they were presented relatively far apart in visual field. However, this type of response
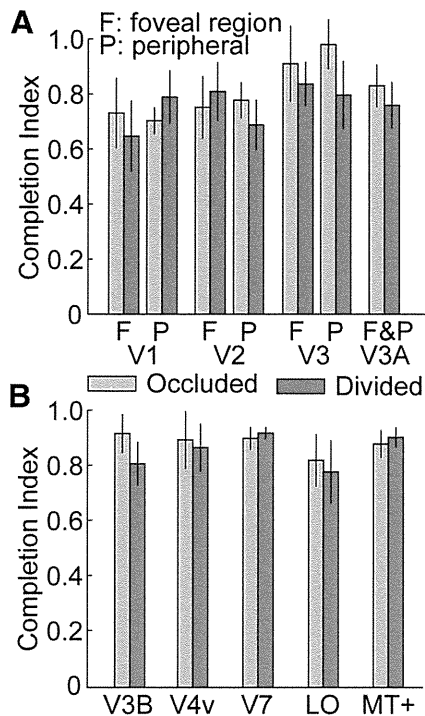
Figure 9. Completion indices in the outer regions of retinotopic ROIs and higher regions. *A,* Completion indices in the foveal/peripheral outer regions of the target subregions that retino-topically represent the occluder positions. *B,* The indices in coarsely retinotopic higher visual areas. Error bars, SEM.

modulation to the visible portions of the partially occluded target was not observed in this study. Thus, neural modulation by amodal completion in early visual areas appears to operate in such a fashion as to simply bridge the occluded portion.

No significant difference was observed in coarsely retinotopic higher areas (paired *t* test; $p > 0.10$ for all the ROIs; Fig. 9B). This was probably due to limitations of spatial resolution of fMRI, since BOLD spread is larger in higher visual areas, and to insensitivity of our retinotopy-based analysis for coarsely retinotopic higher areas. Alternatively, it is also possible that, with our stimulus configuration, these areas do not contribute to topographic completion of an occluded surface, or that they realize more global completion even for a divided target based on synchronized movements of visual elements or visual imagery. Further analysis, though it is beyond the scope of this study, is needed to reach clear conclusions.

Effects of temporally preceding cognitive context

The final question we attempted to answer is whether the observed neural representation of an occluded object is modulated by temporally preceding cognitive context, such as knowledge of the complete appearance of an object obtained before partial occlusion. Answering this question is important for understanding how the brain uses not only spatial (low-level image features) but also cognitive context in maintaining object consistency.

To this end, we presented the divided target rotation with two stable opaque occluders (Fig. 10A, nonoccluded condition). In this configuration, when the target rotated and passed through one of the occluders, the spatial alignments at the overlapping edges such as T-shaped junctions strongly enhanced the impression of occlusion, as in the occluded condition. However, an observer seeing the sequence of target rotation knows that the

target is divided, and not occluded by the occluder. Temporally preceding cognitive experience thus reduces the impression of target occlusion. We explored how this type of incongruent stimulus configuration modulates early topographic responses. We focused on the activity in V1 and V2 in the analysis because reliable amodal completion-related activity was observed only in these areas (Fig. 8A–C).

Responses to this incongruent stimulus mapped on the inflated cortical surface suggested that the temporally preceding cognitive context appeared to have suppressive effect on completion-related activity (Fig. 10B), suggesting in turn that, contrary to common belief, early completion-related activity is not necessarily due to low-level image features alone.

Interestingly, further voxel-by-voxel detailed analyses showed that the response suppression by the cognitive context was only observed in the earliest visual area, V1. Voxel-by-voxel phase scatter-plots revealed that response profiles for the nonoccluded condition were similar with those of the transparent in V1 and V2, whereas periodic powers of the nonoccluded were lower than the transparent especially in V1 (Fig. 10C). Relative periodicity analysis (the occluded periodicity was divided by the nonoccluded one) revealed that neural responses in V1 was indeed suppressed by the temporally preceding cognitive context (*t* test whether the relative periodicity is >1; $t_{(7)} = 2.75, p = 0.014$ in V1; $t_{(7)} = 2.19, p = 0.032$ in V2, only V1 was significant at $p < 0.05$ level after Bonferroni correction; Fig. 10D). Furthermore, one-way (visual polar angles, six locations) repeated-measures ANOVA revealed no main effect of visual field location ($F_{(5,30)} = 0.30, p = 0.91$ in V1; $F_{(5,35)} = 1.80, p = 0.14$ in V2), indicating that cognitive suppression was observed even after the target was completely behind the occluder, and that the results obtained were not due to local differences in stimuli when the target entered the occluder (Fig. 10E).

To assess the cognitive effect in a more conservative fashion, we calculated completion indices. The indices for the nonoccluded condition were significantly lower than those for the occluded only in V1 but not in V2 (paired *t* test, $t_{(7)} = 9.85, p = 0.000012$ in V1; $t_{(7)} = 0.19, p = 0.43$ in V2, only V1 was significant at $p < 0.05$ level after Bonferroni correction; Fig. 11A). These profiles in V1 and V2 were robust even when the ROIs were strictly confined within the center positions of the occluders (but statistically marginally significant in this case; paired *t* test, $t_{(7)} = 1.52, p = 0.086$; Fig. 11B). Furthermore, when we analyzed data after excluding any voxels that showed $p < 0.05$ for the divided stimulus, the index for the nonoccluded was again statistically lower than the occluded only in V1 but not in V2 (paired *t* test; $t_{(7)} = 3.21, p = 0.014$ in V1; $t_{(7)} = 0.50, p = 0.63$ in V2; only V1 was significant at $p < 0.05$ level after Bonferroni correction; Fig. 11C). These results may suggest that observer's prior knowledge does modulate early neural completion-related activity. For indices in foveal/peripheral regions of ROIs, no significant difference was observed between conditions (paired *t* test, $p > 0.09$ for all regions; Fig. 11D).

Interestingly, the higher cognitive effect was found to be stronger in V1 than V2, suggesting that these areas play different roles in solving the occlusion problem; V1 might participate in feedback loops for integrating cognitive context, whereas V2 might be essential for analysis of local image features, such as T-shaped junctions. This discrepancy of response profiles between V1 and V2, together with a case study of patient LG (Gilaie-Dotan et al., 2009), gives a new insight into cortical occlusion representation. LG has abnormal function of intermediate visual areas (V2–V4) whereas fMRI response patterns in V1 and LOC
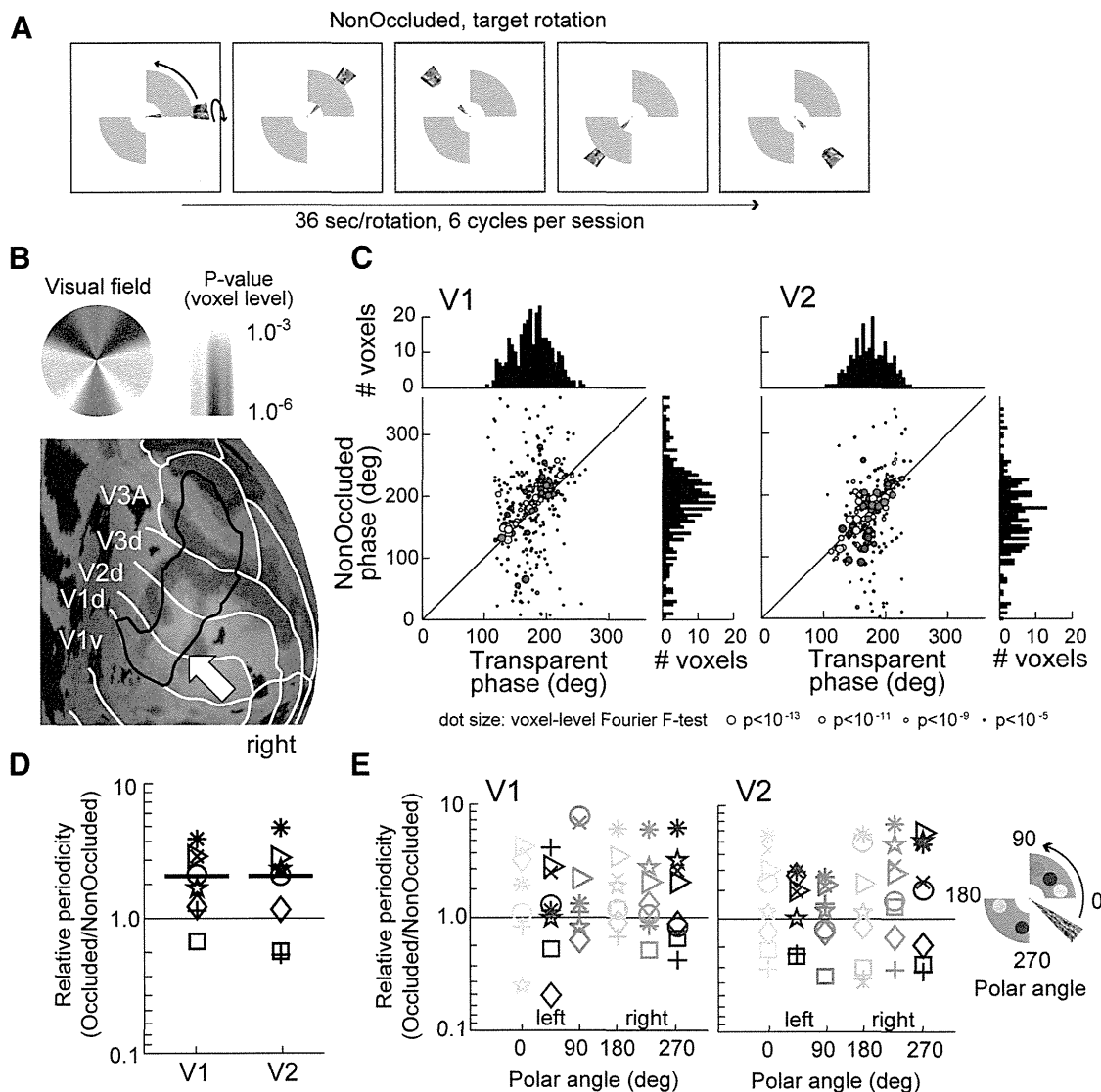
Figure 10. Effect of temporally preceding cognitive context on amodal completion-related activity. *A*, Schematic view of the nonoccluded stimulus presentation. In this configuration, the divided target rotated so as to overlap two stable occluders. Therefore, an observer knows that the target is not occluded but divided, whereas spatial image features such as T-junctions promotes amodal completion percept. *B*, Cortical activity in response to the nonoccluded stimulus. *C*, Voxelwise phase scatter plots, transparent versus nonoccluded. Each color represents a single participant. Dot sizes represent magnitudes of Fourier *F*-statistics (and the corresponding statistical *p* values). For details, see Figure 5. *D*, Relative periodicities (occluded/nonoccluded) exhibited significant decrease in V1 due to temporally preceding cognitive context. Each shape represents a single participant. *E*, Comparison of relative periodic responses (occluded/nonoccluded) in V1 and V2 along polar angle representation. Each color represents the corresponding visual field location. Each shape represents a single participant.

seem to be normal. He has severe deficits with integrating collinear contours and recognizing occluded objects. From this point of view, cortical representations in intermediate levels of the visual processing as well as V1 may be required to solve visual occlusion. Alternatively, it may be possible that relatively local feedback signals (e.g., from V2–V4 to V1) as well as modulations from higher visual areas (see Discussion) would be essential for object completion. It would be worth investigating whether patients with lesions in V2–V4 show topographic neural representation of occlusion in V1.

Attention control experiment and control analyses

The observed topographic activity corresponding to the occluded portion may be alternatively explained as a result of object-based attention to the target (Somers et al., 1999; Mitchell et al., 2004; Lee and Vecera, 2005; Bressler and Silver, 2010; Pratte et al., 2013), or spatial attention to the locations of the occluders (Sa-

saki et al., 2001; Slotnick et al., 2005). To rule out these possibilities, we conducted an additional control experiment using the same stimuli (occluded and divided conditions only) and presentation procedures but with interposing a more difficult attention-demanding task on the central fixation point (see Materials and Methods). As shown in Figure 12, when controlling observer's attention more strictly, the absolute values of the completion index were lower than those of the main experiment (Fig. 8). However, we could still find the significantly higher completion indices for the occluded than the divided in V1/V2 and also in V3 in this experiment (Fig. 12A; paired *t* test; $t_{(7)} = 5.70, p = 0.00073$ in V1; $t_{(7)} = 8.00, p = 0.000091$ in V2; $t_{(7)} = 3.88, p = 0.0061$ in V3; all ROIs were significant at $p < 0.05$ level after Bonferroni correction). In the foveal/peripheral regions of ROIs, no significant difference was observed (Fig. 12B; paired *t* test, $p > 0.10$, for all regions).

It is also possible that attention may work as a noise filter, selectively reducing irrelevant signals at the attended location
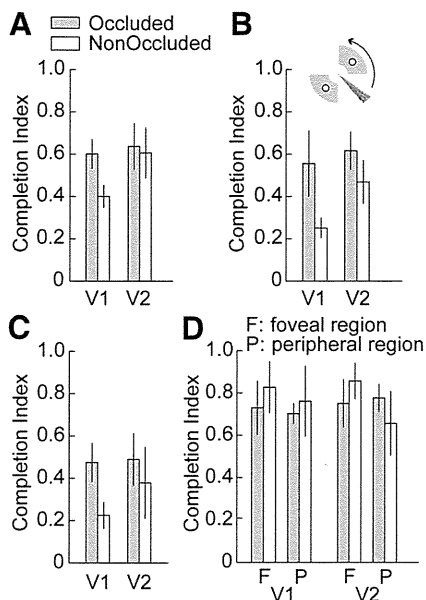
Figure 11. Completion indices for the occluded and nonoccluded stimuli. *A,* Comparison of completion indices for the occluded and nonoccluded stimuli in V1 and V2. *B,* Completion indices after the ROIs in V1/V2 were restricted to regions corresponding to the middle portion of the occluders. *C,* Completion indices in V1 and V2 after excluding all voxels that showed responses at $p < 0.05$ level for the divided condition to minimize BOLD spread effects in the results. The voxel exclusion was done in each scanning day and each participant separately. *D,* Completion indices in the foveal/peripheral regions of ROIs. Error bars, SEM.
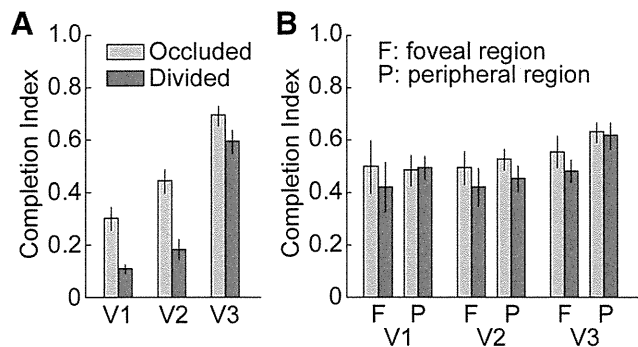

Figure 12. Completion indices for the occluded and divided stimuli with a more attention-demanding task. *A,* Completion indices in retinotopic ROIs under an additional attention control experiment. Although the overall index values decreased in this experiment, higher indices for the occluded over the divided condition were again observed. *B,* Completion indices in the foveal/peripheral regions of ROIs. Error bars, SEM.

(Dosher and Lu, 2000; Lu et al., 2002; Pratte et al., 2013). If this happened in our experiment and the overall noise level for the occluded condition was reduced over the divided and nonoccluded conditions, it would result in apparent response enhancement even without any occlusion completion-related signals. This is because Fourier *F*-statistics were computed by dividing the squared amplitude of power at the target rotation frequency (1/36 Hz) with the sum of squared amplitudes at the other frequencies. To rule out this possibility, we compared noise levels (averages of powers at all the frequencies other than the target rotation frequency and higher harmonics) of time series in each ROI across stimulus conditions in the main experiment. As shown in Figure 13, the results confirmed that the completion-like responses we found in the main experiment cannot be simply explained by overall noise reductions for the occluded condition.

We could not observe any noise level differences across stimulus conditions both in the ROIs (Fig. 13*A*; repeated-measures ANOVA; ROI × stimulus condition; no main effect of stimulus condition, $F_{(2,14)} = 0.076$, $p = 0.93$) and in their foveal/peripheral outer regions (Fig. 13*B*; repeated-measures ANOVA ROI × inner/outer position × stimulus condition; no main effect of stimulus condition, $F_{(2,14)} = 0.13$, $p = 0.88$). We also compared the noise levels for the fMRI time series obtained with a more attention-demanding task (the occluded and divided conditions only). Again, no difference across stimulus conditions was observed both in the ROIs (Fig. 14*A*; repeated-measures ANOVA; ROI × stimulus condition; no main effect of stimulus condition, $F_{(1,7)} = 0.30$, $p = 0.60$) and in the foveal/peripheral regions (Fig. 14*B*; repeated-measures ANOVA; ROI × inner/outer position × stimulus condition; no main effect of stimulus condition, $F_{(1,7)} = 0.79$, $p = 0.40$).

From these results, we can conclude that the observed V1/V2 activity reflects neural amodal completion of the occluded portion. In addition, for all stimulus conditions, we imposed no attentional task related to target rotation and also no task on the location of the occluders. These facts together indicate that the observed topographic activity found here cannot be simply explained by an attention effect alone. Recent neuroimaging studies have also shown association between activity from early/higher visual areas and behavior performance even when the effects of attention were controlled (Sasaki and Watanabe, 2004; Meng et al., 2005; Ban et al., 2006, 2012). Furthermore, as shown in Figure 9*B*, response boosts could not be observed in the external regions of the occluding portions. Therefore, the observed completion responses would not be due to overall attention differences across stimuli and scanning runs. Finally, the observed discrepancy in the responses between V1 and V2 induced by the nonoccluded stimulus (Fig. 11) should not be attributed to simple attentional boosts in early visual areas.

However, we have to keep in mind that it is never entirely possible to completely rule out visuospatial attentional factors from the results. If we imposed an even more challenging fixation task, the withdrawal of attention from the occluded stimulus could also diminish perceptual grouping or binding of the two visible portions of the stimulus even more, resulting in reducing the topographic representation of the whole object in V1/V2. This is consistent with our finding that the knowledge whether the object is divided or whole is relevant to the completion responses. Therefore, although completion-related activity is actually observed, we cannot possibly make a firm conclusion that the completion is derived by an attention-free mechanism. It is not an exclusive scenario; knowing an object is occluded may guide attention to the occluded portion, or attention to the occluded portion may help bind the fragments into a coherent whole.

Discussion
Utilizing dynamically induced occlusion stimuli and investigating concurrent periodic shifts in fMRI activity in early visual cortex, the present study has provided clear human neuroimaging evidence for topographic representation of visual occlusion. We found that V1 and V2 exhibit topographic moment-to-moment activity corresponding to the occluded portion of an object. Based on detailed voxel-by-voxel analysis, we further demonstrated spatial specificity of neural modulation for an amodal object in multiple visual areas. The results of an additional control experiment confirmed that the observed topographic activity cannot be simply explained by attentional boosts of responses in early visual areas. Furthermore, we demonstrated

that the earliest topographic representation of visual occlusion reflects temporally preceding cognitive context as well as spatial image features.

The present study clearly provides evidence of topographic representation of the occluded portion of an object (unseen but sensed object continuity) in human early visual cortex. The topographic representation we found here is consistent with animal single-unit recordings (Sugita, 1999; Bakin et al., 2000) which revealed that some neurons in V1 or V2 respond to an occluded bar within their receptive fields. In these physiological studies, robust amodal completion-related activity was found only when binocular disparity cues between occluding and occluded objects were provided. The present study additionally demonstrated that a pictorial or monocular depth cue provided by rotation of a target is sufficient to cause topographic responses in human V1 and V2. In addition, the topographic activity observed here may reflect neural processes different from that observed previously (Sugita, 1999). In that study, the latency of response in V1 to an occluded bar suggested that completion is mediated via lateral connections or feedback signals coming from areas very close to V1. However, the present study showed higher cognitive effects on topographic activity (Figs. 10, 11), which cannot be explained by activity in early areas alone.

Our finding is consistent with a recent human fMR-adaptation study (Rauschenberger et al., 2006) in that early visual areas contribute to amodal completion. However, it was still unclear what visual feature was represented in early visual areas. That is, it was possible that the observed completion-like V1 activity in fMR-adaptation might simply encode, without completion, specific local image features (Weigelt et al., 2007), such as T-shaped junctions at the occluding borders, one of the striking cues promoting amodal perception (Anderson et al., 2002; Kellman, 2003). To dissociate these possibilities is important to understand how human brain reconstructs stable and unified visual world from fragmental visual inputs. To this end, here, we took a phase-encoded stimulus presentation and analysis technique widely used in the standard retinotopic studies. The method is robust to obtain activities in early visual areas and enabled us to track dynamic fMRI responses related to the occlusion moment-by-moment. By using the dynamic profile of the signals, we could quantitatively evaluate occlusion-related responses with both spatial and temporal specificities. Further, we took several steps to dissociate the topographic neural representation of the occluded portion from other types of neural processing by investigating neural responses within precisely defined cortical subregions.

The current study has clearly shown that visual occlusion is solved accompanied with early topographic representations corresponding to an occluded part of an object. However, it is still not clear how the occluded region of an object is completed.
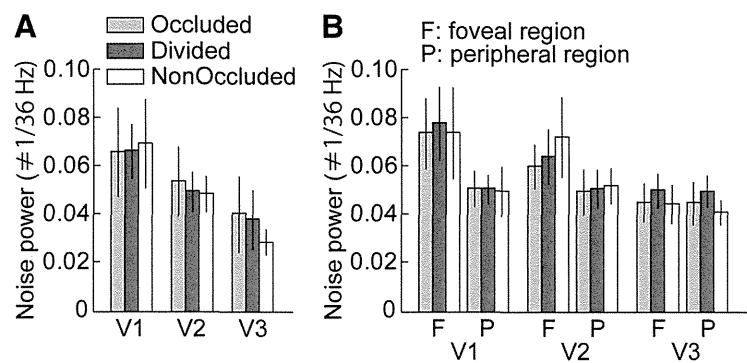

Figure 13. Noise powers in time series for the occluded, divided, and nonoccluded stimuli. *A*, Voxel-averaged noise powers in ROIs for the occluded, divided, and nonoccluded stimuli. Noise powers were computed by averaging Fourier powers other than the target rotation frequency (1/36 Hz) and the higher harmonics. *B*, Voxel-averaged noise powers in the foveal/peripheral regions of ROIs. Error bar, SEM.

Figure 14. Noise powers in time series for the occluded and divided stimuli with a more attention-demanding task. *A*, Voxel-averaged noise powers in ROIs for the occluded and divided stimuli under an additional attention control experiment. Noise powers were computed by averaging Fourier powers other than the target rotation frequency (1/36 Hz) and the higher harmonics. *B*, Voxel-averaged noise powers in the foveal/peripheral regions of ROIs. Error bar, SEM.
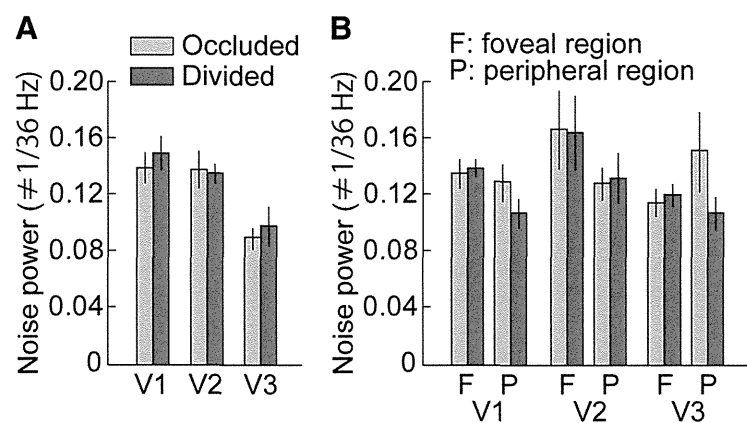
Specifically, we will have two scenarios. One is "surface" completion, in which a whole surface of an occluded object is completed and represented in early visual areas. In contrast, the second possibility is "outline" completion, in which only the outline borders of an occluded portion can be completed. Our findings cannot discriminate these two possibilities. However, we may be able to support outline hypothesis from recent animal and human studies as follows. First, single-cell recording (Von der Heydt et al., 2003) and human fMRI (Cornelissen et al., 2006) have not been able to provide evidence of a topographic representation of a uniform "surface" (but see Sasaki and Watanabe, 2004; Boyaci et al., 2007). In contrast, neurons in early visual areas respond strongly to both real and illusory "borders" (Peterhans and Von der Heydt, 1989; Qiu and von der Heydt, 2007). Based on these findings, we speculate that the observed topographic activity may mainly reflect a representation of completed outline borders, not a uniform surface. Future follow-up study is required to have a conclusive remark on this question.

The observed effect of temporally preceding cognitive context on topographic activity in V1 (Fig. 11) suggests that higher cognitive information and low-level image features are integrated into a unified topographic representation of an occluded object in early visual areas. Some behavioral studies have reported such higher effects on amodal percepts (Joseph and Nakayama, 1999; but see Pratt and Sekuler, 2001; Lee and Vecera, 2005; Palmer et

al., 2006). Our results suggest, though it is counter intuitive, that V1 may be involved in underlying neural substrates of such higher cognitive effects. Then, how are the higher cognitive effects mediated? It is likely that local spatial image features are analyzed through bottom-up cortical pathways. However, it is unlikely that the higher cognitive effects observed here are also mediated exclusively by bottom-up processing. It would be more reasonable to predict that response modulation by cognitive context is responsible for feedback processing from higher areas. This hypothesis may be supported by animal anatomical and physiological studies that demonstrate direct neural connections between higher visual areas and V1 (Hupé et al., 1998; Clavagnier et al., 2004). Lateral occipital (LO) regions that have been extensively studied for their relations with visual object processing would be a most likely candidate for the feedback processing. Although we could not observe any special activity for the non-occluded condition in LO regions in our current experiment, some human fMRI studies actually revealed that occlusion-completed object representations are established at LO processing stage (Kourtzi and Kanwisher, 2001; Lerner et al., 2002, 2004). In addition, one model study suggests that feedback signals from LOC to early visual cortex pertain to the illusory contour percepts (Kogo et al., 2010). The similar model can be applied in the current visual occlusion case. Frontal or prefrontal regions might be the other candidates of higher cognitive modulations. Some studies and models demonstrate frontal effects on human retinotopic visual cortex (Bar, 2003; Ruff et al., 2006; Summerfield et al., 2006).

Our study sheds new light on the roles of V1 activity. Recently, a series of neuroimaging studies have shown that V1 activity and visual awareness are tightly coupled (Polonsky et al., 2000; Pascual-Leone and Walsh, 2001; Ress and Heeger, 2003; Tong, 2003). For example, some studies have demonstrated that V1 responses are correlated with perceptual alternations of binocular rivalry stimuli (Polonsky et al., 2000; Tong, 2003), and that V1 and/or V2 responds to perceptually filled-in (illusory) objects (Peterhans and Von der Heydt, 1989; Meng et al., 2005; Qiu and von der Heydt, 2007). In contrast, in the present study, we showed that V1 robustly responded even to invisible entities, such as occluded portions of objects when participants only "know" that the target is occluded, whereas V1 activity was not directly coupled with "visual" awareness of the occluded portion, because participants did not actually "see" the completed portion of the target. V1 may thus respond not only to what we "see," but also to what we "sense" regardless of whether the target stimulus is visible or not. Future studies will be required to clarify commonalities and differences between neural circuits of filling-in that involves actual percepts and those of amodal completion that does not evoke explicit sensation (Albert, 2007).

What might be the biological advantages of this implicit V1 activity, which reconstructs a topographic representation of an occluded portion of an object? One possibility is utilization of such implicit activity for scene segmentations and object surface reconstructions at the early stage of visual processing. In natural scenes, many surfaces lack counterparts in the retinal image due to occlusion. Thus, reconstructing or completing missing surfaces constitutes a critical intermediate stage in higher object recognition (Nakayama et al., 1995). Implicit activity may reflect such intermediate visual processes, sharing a common neural basis with processing of real and other types of filled-in surfaces. It is also possible that implicit representations may be used more generally; implicit activity may allow other sensory modalities, such as those mediating motor or tactile sensation to interact with invisible objects effectively. For example, implicit representations may aid touching or grasping of occluded portions of objects without difficulty despite lack of vivid visual sensation. Future follow-up studies will be required to have a conclusive remark on the remained mystery of the nature and detailed contents of this implicit V1 activity.

In conclusion, early visual areas represent high-level visual information in the form of a low-level topographic visual world, reflecting not only spatial but also temporal semantic or cognitive contexts. This strategy of unified and topographic cortical processing at the early stage of visual processing is advantageous for recognition of or interaction with objects at the later stages.

References

Albert MK (2007) Mechanisms of modal and amodal interpolation. Psychol Rev 114:445–469. CrossRef Medline

Albright TD, Stoner GR (2002) Contextual influences on visual processing. Annu Rev Neurosci 25:339–379. CrossRef Medline

Anderson BL, Singh M, Fleming RW (2002) The interpolation of object and surface structure. Cogn Psychol 44:148–190. CrossRef Medline

Bakin JS, Nakayama K, Gilbert CD (2000) Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. J Neurosci 20:8188–8198. Medline

Ban H, Yamamoto H (2013) A non–device-specific approach to display characterization based on linear, nonlinear, and hybrid search algorithms. J Vis 13(6):20 1–26. CrossRef Medline

Ban H, Yamamoto H, Fukunaga M, Nakagoshi A, Umeda M, Tanaka C, Ejima Y (2006) Toward a common circle: interhemispheric contextual modulation in human early visual areas. J Neurosci 26:8804–8809. CrossRef Medline

Ban H, Preston TJ, Meeson A, Welchman AE (2012) The integration of motion and disparity cues to depth in dorsal visual cortex. Nat Neurosci 15:636–643. CrossRef Medline

Bar M (2003) A cortical mechanism for triggering top-down facilitation in visual object recognition. J Cogn Neurosci 15:600–609. CrossRef Medline

Bartels A, Zeki S (2000) The architecture of the colour centre in the human visual brain: new results and a review. Eur J Neurosci 12:172–193. CrossRef Medline

Boyaci H, Fang F, Murray SO, Kersten D (2007) Responses to lightness variations in early visual cortex. Curr Biol 17:989–993. CrossRef Medline

Bressler DW, Silver MA (2010) Spatial attention improves reliability of fMRI retinotopic mapping signals in occipital and parietal cortex. Neuroimage 53:526–533. CrossRef Medline

Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. Nat Neurosci 8:1102–1109. CrossRef Medline

Clavagnier S, Falchier A, Kennedy H (2004) Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. Cogn Affect Behav Neurosci 4:117–126. Medline

Cornelissen FW, Wade AR, Vladusich T, Dougherty RF, Wandell BA (2006) No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. J Neurosci 26:3634–3641. CrossRef Medline

Dale AM, Sereno MI (1993) Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci 5:162–176. CrossRef Medline

DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, Miller D, Neitz J (1996) Mapping striate and extrastriate visual areas in human cerebral cortex. Proc Natl Acad Sci U S A 93:2382–2386. CrossRef Medline

Dosher BA, Lu ZL (2000) Noise exclusion in spatial attention. Psychol Sci 11:139–146. CrossRef Medline

Drury HA, Van Essen DC, Corbetta M, Snyder AZ (1999) Surface-based analysis of the human cerebral cortex. In: Brain warping (Toga AW, ed) pp337–363. San Diego: Academic.

Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994) fMRI of human visual cortex. Nature 369:525. CrossRef Medline

Fischl B, Sereno MI, Dale AM (1999) Cortical surface-based analysis: II. Inflation, flattening, and a surface-based coordinate system. Neuroimage 9:195–207. CrossRef Medline

Friedman HS, Zhou H, von der Heydt R (1999) Color filling-in under steady fixation: behavioral demonstration in monkeys and humans. Perception 28:1383–1395. CrossRef Medline

Gilaie-Dotan S, Perry A, Bonneh Y, Malach R, Bentin S (2009) Seeing with profoundly deactivated mid-level visual areas: non-hierarchical functioning in the human visual cortex. Cereb Cortex 19:1687–1703. CrossRef Medline

Hadjikhani N, Liu AK, Dale AM, Cavanagh P, Tootell RB (1998) Retinotopy and color sensitivity in human visual cortical area V8. Nat Neurosci 1:235–241. CrossRef Medline

Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. J Neurosci 22:7195–7205. Medline

Hulme OJ, Zeki S (2007) The sightless view: neural correlates of occluded objects. Cereb Cortex 17:1197–1205. CrossRef Medline

Hupé JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature 394:784–787. CrossRef Medline

Joseph JS, Nakayama K (1999) Amodal representation depends on the object seen before partial occlusion. Vis Res 39:283–292. CrossRef Medline

Kanizsa G (1979) Organization in vision. New York: Praeger.

Kellman PJ (2003) Interpolation processes in the visual perception of objects. Neural Netw 16:915–923. CrossRef Medline

Kogo N, Strecha C, Van Gool L, Wagemans J (2010) Surface construction by a 2-D differentiation-integration process: a neurocomputational model for perceived border ownership, depth, and lightness in Kanizsa figures. Psychol Rev 117:406–439. CrossRef Medline

Komatsu H (2006) The neural mechanisms of perceptual filling-in. Nat Rev Neurosci 7:220–231. CrossRef Medline

Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. Science 293:1506–1509. CrossRef Medline

Kovács G, Vogels R, Orban GA (1995) Selectivity of macaque inferior temporal neurons for partially occluded shapes. J Neurosci 15:1984–1997. Medline

Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. J Neurosci 26:13128–13142. CrossRef Medline

Lee H, Vecera SP (2005) Visual cognition influences early vision: the role of visual short-term memory in amodal completion. Psychol Sci 16:763–768. CrossRef Medline

Lerner Y, Hendler T, Malach R (2002) Object-completion effects in the human lateral occipital complex. Cereb Cortex 12:163–177. CrossRef Medline

Lerner Y, Harel M, Malach R (2004) Rapid completion effects in human high-order visual areas. Neuroimage 21:516–526. CrossRef Medline

Levy I, Hasson U, Avidan G, Hendler T, Malach R (2001) Center-periphery organization of human object areas. Nat Neurosci 4:533–539. CrossRef Medline

Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. In: Proceedings of the 14th annual conference on Computer graphics and interactive techniques. New York: ACM.

Lu ZL, Lesmes LA, Dosher BA (2002) Spatial attention excludes external noise at the target location. J Vis 2(4):4 312–323. CrossRef Medline

Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc Natl Acad Sci U S A 92:8135–8139. CrossRef Medline

Meng M, Remus DA, Tong F (2005) Filling-in of visual phantoms in the human brain. Nat Neurosci 8:1248–1254. CrossRef Medline

Michotte A, Thines G, Grabbe G (1964) Les complements amodaux des structures perceptives. In: Studia psychologica. Louvain, Belgium: Institut de Psychologie de l'Universite de Louvain.

Mitchell JF, Stoner GR, Reynolds JH (2004) Object-based attention determines dominance in binocular rivalry. Nature 429:410–413. CrossRef Medline

Murray MM, Foxe DM, Javitt DC, Foxe JJ (2004) Setting boundaries: brain dynamics of modal and amodal illusory shape completion in humans. J Neurosci 24:6898–6903. CrossRef Medline

Nakayama K, Shimojo S, Ramachandran VS (1990) Transparency: relation to depth, subjective contours, luminance, and neon color spreading. Perception 19:497–513. CrossRef Medline

Nakayama K, He ZJ, Shimojo S (1995) Visual surface representation: a critical link between lower-level and higher-level vision. In: Frontiers in cognitive neuroscience, Ed 2 (Kosslyn S, Osherson DN, eds), pp 1–70. Cambridge, MA: MIT.

Palmer EM, Kellman PJ, Shipley TF (2006) A theory of dynamic occluded and illusory object perception. J Exp Psychol Gen 135:513–541. CrossRef Medline

Pascual-Leone A, Walsh V (2001) Fast backprojections from the motion to the primary visual area necessary for visual awareness. Science 292:510–512. CrossRef Medline

Pessoa L, Thompson E, Noë A (1998) Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. Behav Brain Sci 21:723–802. Medline

Peterhans E, von der Heydt R (1989) Mechanisms of contour perception in monkey visual cortex: II. Contours bridging gaps. J Neurosci 9:1749–1763. Medline

Polonsky A, Blake R, Braun J, Heeger DJ (2000) Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. Nat Neurosci 3:1153–1159. CrossRef Medline

Pratt J, Sekuler AB (2001) The effects of occlusion and past experience on the allocation of object-based attention. Psychon Bull Rev 8:721–727. CrossRef Medline

Pratte MS, Ling S, Swisher JD, Tong F (2013) How attention extracts objects from noise. J Neurophysiol 110:1346–1356. CrossRef Medline

Qiu FT, von der Heydt R (2007) Neural representation of transparent overlay. Nat Neurosci 10:283–284. CrossRef Medline

Rauschenberger R, Liu T, Slotnick SD, Yantis S (2006) Temporally unfolding neural representation of pictorial occlusion. Psychol Sci 17:358–364. CrossRef Medline

Ress D, Heeger DJ (2003) Neuronal correlates of perception in early visual cortex. Nat Neurosci 6:414–420. CrossRef Medline

Ruff CC, Blankenburg F, Bjoertomt O, Bestmann S, Freeman E, Haynes JD, Rees G, Josephs O, Deichmann R, Driver J (2006) Concurrent TMS-fMRI and psychophysics reveal frontal influences on human retinotopic visual cortex. Curr Biol 16:1479–1488. CrossRef Medline

Sasaki Y, Watanabe T (2004) The primary visual cortex filled in color. Proc Natl Acad Sci U S A 101:18251–18256. CrossRef Medline

Sasaki Y, Hadjikhani N, Fischl B, Liu AK, Marrett S, Dale AM, Tootell RB (2001) Local and global attention are mapped retinotopically in human occipital cortex. Proc Natl Acad Sci U S A 98:2077–2082. CrossRef Medline

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268:889–893. CrossRef Medline

Shroeder WJ, Zarge JA, Lorensen WE (1992) Decimation of triangle meshes. In: Proceedings of the 19th annual conference on computer graphics and interactive techniques. New York: ACM.

Slotnick SD, Thompson WL, Kosslyn SM (2005) Visual mental imagery induces retinotopically organized activation of early visual areas. Cereb Cortex 15:1570–1583. CrossRef Medline

Smith AT, Greenlee MW, Singh KD, Kraemer FM, Hennig J (1998) The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). J Neurosci 18:3816–3830. Medline

Somers DC, Dale AM, Seiffert AE, Tootell RB (1999) Functional MRI reveals spatially specific attentional modulations in human primary visual cortex. Proc Natl Acad Sci U S A 96:1663–1668. CrossRef Medline

Sugita Y (1999) Grouping of image fragments in primary visual cortex. Nature 401:269–272. Medline

Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J (2006) Predictive codes for forthcoming perception in the frontal cortex. Science 314:1311–1314. CrossRef Medline

Teo PC, Sapiro, G. Wandell A (1997) Creating connected representations of cortical gray matter for functional MRI visualization. IEEE Trans Med Imaging16:852–863. CrossRef Medline

Tong F (2003) Primary visual cortex and visual awareness. Nat Rev Neurosci 4:219–229. CrossRef Medline

Tootell RB, Mendola JD, Hadjikhani NK, Ledden PJ, Liu AK, Reppas JB, Sereno MI, Dale AM (1997) Functional analysis of V3A and related areas in human visul cortex. J Neurosci 17:7060–7078. Medline

Von der Heydt R, Friedman H, Zhou H (2003) Searching for the neural mechanisms of color filling-in. In: Filling-in: from perceptual completion to skill learning (Pessoa L, de Weerd P, eds), pp 106–127. New York: Oxford UP.

Wandell BA, Brewer AA, Dougherty RF (2005) Visual field map clusters in human cortex. Philos Trans R Soc Lond B Biol Sci 360:693–707. CrossRef Medline

Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. Neuron 56:366–383. CrossRef Medline

Weigelt S, Singer W, Muckli L (2007) Separate cortical stages in amodal completion revealed by functional magnetic resonance adaptation. BMC Neurosci 8:70. CrossRef Medline

Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC (1998) Automated image registration: I. General methods and intrasubject, intramodality validation. J Comput Assist Tomogr 22:139–152. CrossRef Medline

Yamamoto H, Ban H, Fukunaga M, Umeda M, Tanaka C, Ejima Y (2008) Large- and small-scale functional organization of visual field representation in the human visual cortex. In: Visual Cortex: New Research (Portocello TA, Velloti RB, eds), pp195–226. New York: Nova Science.

Yamamoto H, Fukunaga M, Takahashi S, Mano H, Tanaka C, Umeda M, Ejima Y (2012) Inconsistency and uncertainty of the human visual area loci following surface-based registration: probability and entropy maps. Hum Brain Mapp 33:121–129. CrossRef Medline

Yin C, Shimojo S, Moore C, Engel SA (2002) Dynamic shape integration in extrastriate cortex. Curr Biol 12:1379–1385. CrossRef Medline

Zeki S, Watso JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991) A direct demonstration of functional specialization in human visual cortex. J Neurosci 11:641–649. Medline

Behavioral/Cognitive

Dynamic Neural Network Reorganization Associated with Second Language Vocabulary Acquisition: A Multimodal Imaging Study

Chihiro Hosoda,[1,2,3,4] Kanji Tanaka,[5,6] Tadashi Nariai,[3] Manabu Honda,[1,2] and Takashi Hanakawa[1,2,7]

[1]Department of Functional Brain Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan, [2]Department of Advanced Neuroimaging, Integrative Brain Imaging Center, National Center of Neurology and Psychiatry, Tokyo 187-8551, Japan, [3]Department of Neurosurgery, Tokyo Medical and Dental University, Tokyo 113-0034, Japan, [4]Department of Motor Control and Rehabilitation, ATR Computational Neuroscience Laboratories, Kyoto, 619-0288, Japan, [5]Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, 113-8654, Japan [6]Japan Society for the Promotion of Science, Tokyo, 102-0083, Japan, and [7]PRESTO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan

It remains unsettled whether human language relies exclusively on innately privileged brain structure in the left hemisphere or is more flexibly shaped through experiences, which induce neuroplastic changes in potentially relevant neural circuits. Here we show that learning of second language (L2) vocabulary and its cessation can induce bidirectional changes in the mirror-reverse of the traditional language areas. A cross-sectional study identified that gray matter volume in the inferior frontal gyrus pars opercularis (IFGop) and connectivity of the IFGop with the caudate nucleus and the superior temporal gyrus/supramarginal (STG/SMG), predominantly in the right hemisphere, were positively correlated with L2 vocabulary competence. We then implemented a cohort study involving 16 weeks of L2 training in university students. Brain structure before training did not predict the later gain in L2 ability. However, training intervention did increase IFGop volume and reorganization of white matter including the IFGop-caudate and IFGop-STG/SMG pathways in the right hemisphere. These "positive" plastic changes were correlated with the gain in L2 ability in the trained group but were not observed in the control group. We propose that the right hemispheric network can be reorganized into language-related areas through use-dependent plasticity in young adults, reflecting a repertoire of flexible reorganization of the neural substrates responding to linguistic experiences.

Introduction

Language relies on both innate ability and experiences (Kuhl, 2010; Hsu et al., 2011), making understanding of experience-dependent shaping of language systems an essential theme for language neuroscience. The neural mechanisms of experienced-induced first language (L1) development are difficult to study since L1 is acquired through infancy-childhood. Understanding of second language (L2) learning is therefore important because it should also invoke basic mechanisms of language acquisition.

Vocabulary is important for proficiency in any language (Coady and Huckin, 1997; Grogan et al., 2012), but neural mecha-

nisms of L2 vocabulary learning remain unclear. L2 vocabulary learning is achieved through interactions with the already existing L1 lexicon in adults (Ellis et al., 1999). For example, L2 stimuli spontaneously activate the L1 lexicon (Thierry and Wu, 2007), and individuals with high L1 vocabulary competence have an advantage in L2 vocabulary learning (Meschyan, 2002). However, despite these interactions, the L1 and L2 lexicons may be differentially structured; it has been proposed that the L2 lexicon is more phonologically organized compared with the L1 lexicon strongly connected with the semantic system (Laufer, 1989). Hence, L2 vocabulary leaning in adults would induce reorganization predominantly in the phonological system nested with the semantic system. These features may make L2 vocabulary learning in adults a unique paradigm for exploring repertoires of dynamic neural reorganization of the language system, especially in the case of linguistically distant languages such as English and Japanese (Chiswick and Miller, 2005).

L2 learning may accompany macroscopic structural alterations of the language system. Recent neuroimaging studies have revealed plastic changes of gray matter (GM) during learning of various cognitive and motor abilities (Draganski et al., 2004; Mechelli et al., 2004). The analysis of fractional anisotropy (FA) computed from diffusion-weighted magnetic resonance imaging

Table 1. Means and SDs of subject demographics, L2 proficiency tests, IQ, and personality assessments (NEO-FFI)

	Cross-sectional study (n = 137)	Cohort study (pretraining data)	
		TG (n = 24)	CG (n = 20)
Age	21.8 ± 16.0	20.1 ± 1.8	20.5 ± 1.7
Sex (male: female)	66:71	10:14	10:10
TOEIC (200)	NA	102.5 ± 2.6	100.8 ± 2.3
EVT (100)	20.6 ± 16.0	11.3 ± 10.9	12.9 ± 12.1
NART (100)	41.8 ± 6.2	35.6 ± 4.5	39.8 ± 4.0
WAIS-TIQ	NA	113.0 ± 2.2	116.5 ± 1.8
WAIS-VIQ	NA	101.0 ± 2.6	104.0 ± 2.1
WAIS-PIQ	NA	103.5 ± 2.5	101.0 ± 2.3
NEO-FFI neuroticism	NA	25.4 ± 1.4	26.4 ± 1.4
NEO-FFI extraversion	NA	29.8 ± 1.4	30.6 ± 1.3
NEO-FFI openness	NA	32.2 ± 1.3	32.3 ± 1.3
NEO-FFI agreeableness	NA	33.9 ± 1.0	34.1 ± 1.1
NEO-FFI conscientiousness	NA	31.5 ± 1.3	30.7 ± 1.2

In the cohort study, no significant differences were found between the TG and CG for all measurement types. NA, not available; WAIS-VIQ, Wechsler Adult Intelligence Scale Verbal IQ; WAIS-TIQ, Wechsler Adult Intelligence Scale Total IQ; WAIS-PIQ, Wechsler Adult Intelligence Scale Performance IQ; NEO-FFI, NEO Five-Factor Inventory.

(DWI) can index organization of white matter (WM) microstructure, and allows us to visualize plastic changes of WM (Johansen-Berg, 2012; Zatorre et al., 2012). Very recently, a few L2 learning studies have shown GM (Mårtensson et al., 2012) and WM (Schlegel et al., 2012) changes in the traditional language network in the left hemisphere.

Previous neuroimaging studies have, however, demonstrated roles of the right hemisphere in language ability (Vingerhoets et al., 2003; Videsott et al., 2010; Vigneau et al., 2011; Van Ettinger-Veenstra et al., 2012) and in L2 control (Hernandez et al., 2001; Hosoda et al., 2012). One of these found correlations between right prefrontal activity and L2 vocabulary competence (Hosoda et al., 2012). To explore neural plasticity associated with L2 vocabulary learning, we first explored the neural underpinnings correlated with L2 vocabulary competence in a cross-sectional study, revealing GM and WM structures correlated with L2 vocabulary levels predominantly in the right hemisphere. We then ran a cohort study implementing an L2 training program to test whether those L2 vocabulary correlates reflected innate predispositions or plastic changes resulting from learning. We further examined the natural course of the induced structural changes. Here we provide evidence that L2 vocabulary learning induces dynamic reorganization of GM and WM structures outside of the typical language network in adult brains.

Materials and Methods

Subjects. In the cross-sectional study, 137 native Japanese speakers (71 males and 66 females) with a mean age of 24.0 years [SD = 5.3, range 18–42] were enrolled (Table 1). All subjects were university students or graduates whose self-reported English (L2) proficiencies varied from low to very high to the level of Japanese–English bilinguals.

In the cohort study, 67 native Japanese speakers (mean age = 20.1, 31 men and 36 women) were initially placed into a training group (TG). In the TG, 24 of 47 participants completed the planned 4 month training, whereas the remaining 23 participants opted not to continue during the course of training for various reasons. Hence, we studied 24 TG participants (mean age = 20.1, 10 men and 14 women) and an age-matched control group (CG; n = 20; mean age = 20.1, 10 men and 10 women). The participants in the cohort study were university students and recruited from basic level English classes (Table 1). The participants in the cross-sectional study and those in the cohort study did not overlap.

In both studies, the participants had grown up in Japan and started to learn English as their L2 at a mean age of 11.0 years. None had started to learn L2 before 7 years old, and thus all participants were regarded as late L2 learners (Dowens et al., 2010). All the participants were right-handed

as assessed by the Edinburgh Handedness Inventory. All of the participants showed handedness scale of 100 (SD = 0) because we only recruited the perfectly right-handed participants, revealing the contents of handedness inventory in the recruitment information. All were healthy and neurologically intact, with no history of neuropsychiatric disorders, psychotropic medication use, or head injury. All the participants gave written informed consent according to the study protocol approved by the institutional review board (National Center of Neurology and Psychiatry).

Experimental design. In the cross-sectional study, the participants underwent MRI scanning (T1-weighted images and DWI) and assessment for English proficiency using the English Vocabulary Test (EVT) and the National Adult Reading Test (NART). The T1-weighted MRI provided the data for voxel-based morphometry (VBM) analysis of the GM volume, and the DWI allowed for tract-based spatial statistics (TBSS) and probabilistic diffusion-based tractography (PDT).

In the cohort experiment, we adopted the Test of English for International Communication (TOEIC) as a primary measure to evaluate various L2 abilities. The TOEIC assesses English proficiency for use in business, providing a fairly accurate measurement of English capabilities for non-native speakers in listening, reading, and grammar. The EVT and NART were used as adjunctive measures of L2 proficiency. Additionally, the Wechsler Adult Intelligence Scale-3 (WAIS-3) and NEO-Five Factor Inventory (NEO-FFI) were used to confirm homogeneity of basic intelligence and personality traits between the TG and CG.

All participants underwent MRI scanning (T1-weighted MRI and DWI) and an L2 assessment battery before training, hereafter referred to as the "Pre" condition. The L2 training program was developed in-house using Visual Basic. Participants undertook 16 weeks of training, and in each week they learned 60 words or idioms including the meaning, spelling, and pronunciation as well as example sentences indicating their usage. For pronunciation, we referred to the AT&T Natural Voices Text-to-Speech Demo (http://www2.research.att.com/~ttsweb/tts/demo.php). The participants were encouraged to dictate each word, idiom, and sentence 10 times. Each weekend, we distributed a program that included a review test using the 60 words or idioms of the week. Upon completing the full program, the participants were expected to master almost 1000 words and idioms. Participants in the CG received no particular assignment according to previous cohort studies using e-learning (Takeuchi et al., 2010; Mårtensson et al., 2012; Schlegel et al., 2012; Wan et al., 2012; Ghazi Saidi et al., 2013).

The participants in both the TG and CG underwent an L2 assessment battery immediately after the training period ("Post-1"). After completion of the training program, the participants were not further engaged in specific L2 learning programs systemically, and thus the continuation of L2 learning depended upon each participant's choice in personal life. We obtained follow-up behavioral and imaging data from the TG group a year after the end of the training program ("Post-2"). Statistical tests on the behavioral parameters were performed using SPSS 17.0 (IBM). To explore training-induced changes in L2 abilities, we analyzed total TOEIC score and its subsections (listening, reading, and grammar) using a 2-by-2 mixed repeated-measures ANOVA with the time (Pre and Post-1) as a within-subject variable and the group (TG and CG) as a between-subject variable. Next, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of EVT or NART score.

Image data acquisition. MRI data were acquired using a 3 T MRI scanner (Siemens Trio) with an 8-channel phased array receiver coil. High-resolution, 3D, T1-weighted anatomical images were obtained with an MPRAGE sequence designed as follows: repetition time (TR) = 2000 ms, echo time (TE) = 4.4 ms, inversion time (TI) = 990 ms, flip angle = 80°, matrix size = 192 × 176, field of view (FOV) = 192 × 176 mm and 1 mm^3 isotropic voxels. We also acquired whole-brain DWI as follows: TR = 7900 ms, TE = 80 ms, 65 slices, flip angle = 90°, matrix size = 96 × 96, FOV = 192 × 192 mm, 2 × 2 × 2 mm^3 isotropic voxels, 81 volumes with diffusion weighting (b value = 700 s/mm^2) for different motion probing gradient directions and 9 volumes without diffusion weighting (b = 0 s/mm^2). Field-map images were acquired in the same scanning space as the DWI (TE1 = 5.19 ms; TE2 = 7.65 ms). All image data were

converted into the Neuroimaging Informatics Technology Initiative format before further processing.

Image data analysis: GM-VBM. The high-resolution 3D T1-weighted images were subjected to a VBM analysis using VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm.html) implemented in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). The preprocessing steps were as follows. (1) Each image was segmented into GM, WM, and CSF images in the native image space. (2) The diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL) registration method was used to create a study-specific mean GM image template by using the aligned images from all the subjects to improve intersubject registration (Ashburner, 2007). Individual's GM images were registered to the study-specific mean GM template. (4) The registered GM images were further transformed into the Montreal Neurological Institute (MNI) space. (5) These normalized GM images were smoothed using a Gaussian kernel of 12 mm full-width at half-maximum.

In the cross-sectional study, we investigated the correlations of EVT or NART scores with regional GM volume. A multiple regression design was used using EVT or NART score as an independent variable and sex, duration of L2 learning, and age of acquisition as nuisance variables. We set the voxelwise significance level at $p < 0.05$ corrected for multiple comparisons in terms of the familywise error (FWE) rate. After identifying the inferior fontal gyrus pars opercularis (IFGop), superior temporal gyrus (STG), and caudate nucleus as the correlates of L2 vocabulary, we further investigated which hemisphere was more relevant. We computed the laterality index (LI) of GM volume correlated with L2 vocabulary competence (EVT), by using mean β-values computed from 5 mm spherical volumes of interest (VOIs) in IFGop ($x, y, z = \pm 40, 9, 24$), STG/supramarginal gyrus (SMG; $x, y, z = \pm 57, -36, 6$), and caudate nucleus as follows ($x, y, z = \pm 6, 8, 8$). All of these VOIs were defined in the left hemisphere, and thus mirror-reversed VOIs were created for the right hemisphere.

$$LI = (\text{mean}\beta_{\text{LEFT\_VOI}} - \text{mean}\beta_{\text{RIGHT\_VOI}})/$$

$$(\text{mean}\beta_{\text{LEFT\_VOI}} + \text{mean}\beta_{\text{RIGHT\_VOI}})$$

The selection of VOI (IFGop, STG, and caudate nucleus) was based on the results from previous studies (Crinion et al., 2006; Saur et al., 2008). The LI ranges from -1 (completely lateralized to the right) to $+1$ (completely lateralized to the left). Individuals with an LI of $> +0.4$ or < -0.4 were categorized into the left or right hemisphere dominant group, respectively, while those with LI between -0.4 and $+0.4$ were placed into the bilateral representation group (Briellmann et al., 2003).

In the cohort study, we first tested the possibility that GM volume before L2 training could predict the degree of L2 ability improvement after training. Positive results from this analysis would support the hypothesis that particular brain regions are reserved for L2 vocabulary learning. If such were the case, people with particularly developed structures somewhere in the brain might show better L2 learning ability. To test this hypothesis, a correlation analysis was performed using the MRI data of the Pre condition and the training-related improvements in L2 ability (differences in TOEIC scores between the Pre and Post-1 conditions). Next, to explore training-induced changes in GM, we conducted a 2-by-2 mixed repeated-measures ANOVA with time (Pre and Post-1) as a within-subject variable and group (TG and CG) as a between-subject variable. The regions showing significant time-by-group interaction ($p < 0.05$, FWE corrected) were explored. Third, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of GM volume in right IFGop. This analysis tested if individual differences in brain plastic changes accounted for individual differences in L2 performance improvement after training.

Image data analysis: DWI-TBSS. Data preprocessing and analysis of DWI were performed using Oxford Centre for Functional MRI of the Brain (FMRIB)'s software library (FSL 4.1, UK; http://www.fmrib.ox.ac.uk/fsl/]. All DWI were registered to the b0 images. Nonlinear image distortions due to magnetic field (b0) inhomogeneity were corrected based on the field map images, using FUGUE in the FMRIB software library. The registered images were skull stripped using the Brain Extraction Tool. FA maps were calculated using the FMRIB's Diffusion Tool-

box (FDT) v2.0 (Smith et al., 2004). After calculation of the FA map for each subject, we implemented a voxelwise statistical analysis of the FA data using TBSS v1.2 (Smith et al., 2006). In brief, TBSS was performed as follows: (1) alignment of the FA images from all subjects to a template that was arbitrarily selected from those FA images, (2) transformation of all the aligned FA images into MNI space using affine registration to remove the effect of cross-subject spatial variability (Kerns et al., 2004), (3) creation of a mean FA image and FA skeleton images corresponding to the center of the WM using a threshold of FA ≥ 0.20 (Smith et al., 2006), (4) projection of the individuals' FAs onto the mean FA skeleton, and (5) voxelwise cross-subject statistical analyses.

As water molecules move faster along the direction of WM fibers than in the direction perpendicular to them, the direction and size of water diffusion in the brain can index structural organization of WM fibers (Le Bihan and Johansen-Berg, 2012; Zatorre et al., 2012). DWI can measure water diffusion noninvasively *in vivo*. FA is a parameter computed from DWI according to tensor-based modeling of water diffusion, and represents the degree of directional bias of water diffusion. FA is sensitive to size and density of axons, degrees of myelination, and the coherence of organization of fibers within a voxel (Beaulieu, 2002; Alexander et al., 2007; Zatorre et al., 2012) In the cross-sectional study, correlations of L2 proficiency (EVT or NART) with FA values as an index of WM organization were tested using Randomize v2.1 in FSL. The statistical threshold was set at $p < 0.05$ (FWE corrected), and the threshold-free cluster enhancement method was used to define the clusters.

In the cohort study, similar to the VBM analysis, we first tested the possibility that structural organization of WM before L2 training could predict the degree of L2 ability improvement after training. Recent literature indicates that training-induced changes of FA can capture plastic reorganization of WM (Johansen-Berg, 2007, 2012; Flöel et al., 2009; Scholz et al., 2009; Johansen-Berg et al., 2012; Zatorre et al., 2012). Activity-dependent myelo-modulation is a potential mechanism by which WM is carved by experiences. Such structural WM changes, including production of myelin basic protein, are correlated with changes in FA (Blumenfeld-Katzir et al., 2011). Changes in myelination may alter conduction velocity and synchronization of signal transmission across the remote areas (Fields, 2005), relating probably to behavioral changes. To capture such learning-induced reorganization of WM, we conducted a 2-by-2 mixed repeated-measures ANOVA with time as a within-subject variable and the group as a between-subject variable ($p < 0.05$, FWE corrected), yielding FA changes specific to the L2 training program. A correlation analysis was performed using FA images (Pre) and the changes in TOEIC scores from the Pre to Post-1 conditions.

Image analysis: multifiber PDT. We conducted PDT using FDT (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT). We performed tractography by tracing pathways through the estimates of diffusion directions. A multifiber model was used to handle the issue of fiber crossing. In each voxel, two fiber directions were modeled, and a probability distribution function (pdf) of diffusion parameters was estimated. We then proceeded with PDT, drawing multiple (5000) streamline samples based on the pdf and estimated the distribution of connections from each seed to target voxels. The generated pathways were volumes in which values at each voxel represented the number of samples passing through that voxel. The voxel value corresponded to the probability of the target voxel connecting to the seed voxel. To remove spurious nonsignificant connections, pathways in individual subjects were thresholded to include only voxels that had at least 50 samples passing through them (out of 5000 samples drawn from each seed voxel).

Previous studies have pointed to the role of the caudate nucleus in L2 (Crinion et al., 2006; Abutalebi et al., 2008; Grogan et al., 2009; Hosoda et al., 2012; Mårtensson et al., 2012). Moreover, the present TBSS result suggested a correlation between L2 ability and WM situated between the IFG and the caudate nucleus and temporoparietal junction (STG/SMG). In the cross-sectional study, therefore, a correlation analysis was performed to investigate the correlation between an L2 proficiency index (EVT or NART score) and the WM connectivity parameter computed from the bilateral IFGop-caudate and IFGop-STG/SMG corresponding to the dorsal pathway (Rilling et al., 2008; Saur et al., 2008). The whole caudate nucleus was manually defined for a target mask in each hemi-
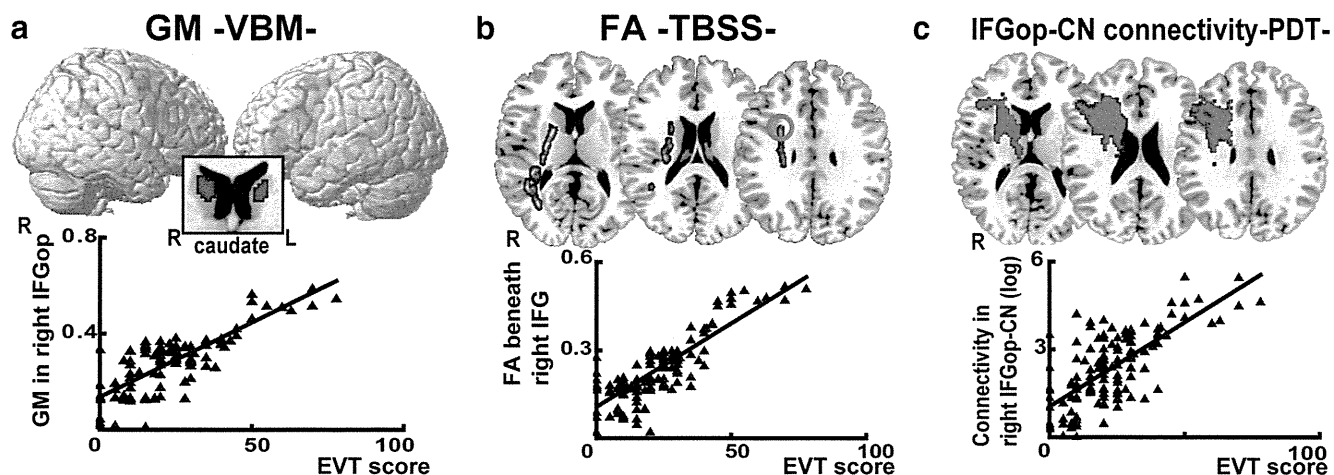
Figure 1. Results from the cross-sectional study. *a*, 3D rendered images showing the correlation between EVT score and GM volume from the VBM analysis. GM volume for the bilateral IFGop, STG/SMG, and caudate nucleus (CN) correlated with EVT score, reflecting differential levels of L2 proficiency across participants (*p* < 0.05 FWE corrected). The plots show the correlation between EVT and GM volume in the right IFGop (green circle). *b*, The TBSS analysis displays voxels with a positive correlation between EVT score and fractional anisotropy reflecting integration of WM structure. Significant correlation was found in the subcortical region beneath the IFGop and the AF. The plots show the correlation between EVT scores and FA values in the WM beneath the right IFGop (green circle). *c*, The PDT analysis shows fiber connections between the right IFGop and the CN in all participants. The EVT score was significantly correlated with connectivity of the IFGop-caudate head (plot). R, right; L, left.

sphere according to each subject's 3D T1-weighted image (Peltier et al., 2011). A spherical VOI with a 5 mm radius was used as a seed mask for the IFGop. The dorsal pathway corresponds to fibers connecting IFGop and STG, and includes arcuate fasciculus (AF) and parts of superior longitudinal fasciculus (SLF; Rilling et al., 2008; Saur et al., 2008). According to a previous study (Saur et al., 2008), we implemented 5 mm radius spherical VOIs in the bilateral IFGop and STG/SMG for the dorsal pathway. We performed PDT within each hemisphere: IFGop-caudate nucleus and IFGop-STG/SMG (dorsal pathway) (four tracts in total).

In the cohort study, in addition to the IFGop-caudate and IFGop-STG/SMG pathways, we also assessed correlations between the L2 competence and other long frontoparietal/temporal association tracts that are reported to be relevant to language processing/learning. We considered the ventral language pathways and inferior longitudinal fasciculus (ILF). The ventral pathway connects IFG pars triangularis (IFGtri) and middle temporal gyrus (MTG), passing through the extreme capsule. ILF connects medial temporal lobe with the temporo-occipital junction, and is suggested be relevant to lexical-semantic language processing (Catani and Mesulam, 2008; Rilling et al., 2008; Saur et al., 2008). We implemented 5 mm radius spherical VOIs in the bilateral IFGtri (BA45, $x, y, z = \pm48, 27, 12$) and MTG ($x, y, z = \pm48, -60, 18$) for the ventral pathway. We placed VOIs in the medial temporal lobe ($x, y, z = \pm40, -20, -7$) and temporo-occipital junction ($x, y, z = \pm35 -68, -6$) for ILF according to a previous study (Jou et al., 2011). We performed PDT within each hemisphere: IFGop-caudate nucleus, IFGop-STG/SMG (dorsal pathway), IFGtri-MTG (ventral pathway), and ILF (eight tracts in total). The voxel values indexing probabilistic strength of connectivity across the two regions were averaged within each PDT pathway and then subjected to logarithmic transformation, producing a parameter for statistical analyses (hereafter called the connectivity parameter). Statistical tests on the connectivity parameter computed from PDT were performed using SPSS 17.0 (IBM).

In the cohort study, we performed a 2-by-2 mixed repeated-measures ANOVA with time as a within-subject variable and the group as a between-subject variable. In addition, a correlation analysis was performed to test if the training-related changes in TOEIC score were correlated with those of the connectivity parameter from the eight tracts: bilateral IFGop-caudate pathway, dorsal pathway, ventral pathway, and ILF. Moreover, to test the possibility that connectivity of the specific tracts before intervention could predict the degree of improvement in L2 ability, a correlation analysis was performed between the connectivity parameter of the PDT in the Pre condition and the improvement of L2 proficiency.

Analysis of follow-up behavioral and imaging data. Only in the training group, for which follow-up data were available, TOEIC score, GM volume, FA value, and the strength of connectivity of the eight specific tracts were compared across the Pre, Post-1, and Post-2 conditions. To retrieve the GM volume and FA values of interest, we applied spherical VOIs with a 5 mm radius to the right IFGop and WM beneath the right IFG based on the results from the cross-sectional study. The data were fed into a one-way repeated-measures ANOVA, followed by *post hoc* comparisons with Tukey's HSD test. Finally, we ran a correlation analysis between the GM changes from the Pre to Post-1 or Post-1 to Post-2 and FA changes from the Pre to Post-1 or Post-1 to Post-2. This analysis suggested the existence of different subgroups following the completion of training. Indeed, three participants continued L2 learning of their own motivation after our intervention, while the rest of the participants ($n = 21$) did not. Hence, a subgroup analysis was conducted to compare the changes of TOEIC, GM, or FA from Post-1 to Post-2 between the two groups. Because one of the groups consisted of a small number of participants, we performed this test with a nonparametric, Mann–Whitney *U* test.

Results

Cross-sectional study
Mean EVT and NART scores were 20.6 and 41.8, respectively (Table 1). The large variance of these values supported that L2 proficiency of the participants was variable.

The VBM analysis showed that individuals with more extensive L2 vocabulary (EVT) had significantly larger GM volume in the IFGop corresponding to Brodmann area (BA) 44, caudate nuclei, STG/SMG, and anterior cingulate cortex, all bilaterally (Fig. 1, Table 2). The correlation of the EVT score and GM volume was found more strongly in the right hemisphere for the IFGop (LI = −0.40) and within the range of symmetry for the caudate nucleus (LI = −0.15) and STG (LI = −0.17). The TBSS analysis showed that individuals with richer L2 vocabulary showed higher FA values in the subcortical WM beneath the IFGop (sub-IFGop), ILF, and AF only in the right hemisphere. The PDT identified bilateral IFGop-caudate nucleus and IFGop-STG/SMG (dorsal pathway) in all individuals. A statistical analysis of the connectivity parameters retrieved from the four tracts revealed correlation of the EV score with connectivity of the IFGop-caudate nucleus ($p < 0.001$) and the IFGop-STG/SMG (dorsal pathway) ($p < 0.001$) in the right hemisphere, but not in

Table 2. Significant correlation between EVT score and GM volume ($p < 0.05$ corrected) in the cross-sectional study

Anatomical location	Coordinates			Z-value	P corrected
	x	y	z		
Right IGFtri (BA44)	36	11	28	7.84	0.00
Right caudate nucleus	12	8	19	5.92	0.00
Right STG/SMG	56	−31	21	5.42	0.00
Right MTG	49	−20	0	4.08	0.00
Left IFG (BA44, 45)	−38	16	30	6.69	0.00
Left superior frontal gyrus	−22	35	45	5.45	0.00
Left caudate nucleus	−4	16	−2	5.26	0.00
Left STG/SMG	−48	−50	26	4.92	0.00

The coordinates (x, y, z) indicate local maxima in each brain region according to the MNI template.

the left hemisphere (IFGop-caudate, $p = 0.09$; dorsal pathway, $p = 0.11$; Fig. 1).

NART score did not correlate with GM volume (VBM analysis) or FA value (TBSS analysis) at the predetermined statistical threshold. When a lenient threshold (uncorrected $p < 0.001$) was applied to the VBM analysis, we detected a trend toward correlation between NART score and GM volume in the middle temporal and postcentral gyri in the left hemisphere.

Cohort study: behavioral data

Baseline profiles did not differ between the TG and CG participants in terms of age, sex, IQs, personality traits, or L2 ability (Table 1).

The total learning time, recorded on log files, was 45.5 h (SD = 3.6, range 25.1 − 64.6) across the whole training period. Just after L2 training (Post-1), the TG showed a 29 ± 21.3% (mean ± SD) improvement in total TOEIC score, whereas the CG group showed no change (0 ± 8.1%; Fig. 2a), revealing the significant effect of L2 training ($F_{(3,82)} = 27.15$, $p < 0.001$ by mixed repeated-measures ANOVA; Table 3). Specifically, improvement was found in the listening ($F_{(3,82)} = 17.64$, $p < 0.001$) and reading ($F_{(3,82)} = 10.15$, $p = 0.001$) sections of TOEIC, but not in the grammar section ($F_{(3,82)} = 0.50$, $p = 0.42$; mixed repeated-measures ANOVA). This finding is reasonable because the training program included items for vocabulary and listening (i.e., pronunciation), but not for grammar. The changes in TOEIC score paralleled those of the EVT score ($r = 0.41$, $p = 0.03$), whereas no correlation was found with those of the NART score ($r = 0.15$, $p = 0.53$).

Follow-up L2 proficiency tests were obtained in the TG participants a year after the training program. One-way repeated-measures ANOVA showed significant differences in TOEIC score across the three time points ($F_{(1.8, 38.6)} = 24.96$, $p < 0.001$ by repeated-measures ANOVA with sphericity correction). Post hoc comparisons (Tukey's HSD test) demonstrated significant increases in score from the Pre to Post-1 ($p < 0.001$) conditions and significant decreases from Post-1 to Post-2 ($p = 0.02$), resulting in no difference in score between Pre and Post-2.

Cohort study: imaging data

Before the training (Pre), no significant differences in brain structure were detected in the GM-VBM, WM-TBSS, and connectivity of the eight specific tracts (bilateral IFGop-caudate, dorsal pathway, ventral pathway, and ILF) between the TG and the CG, supporting the homogeneity of the two groups. Further, we examined whether brain architecture at the Pre stage could predict L2 ability improvement gains after the training program in the TG. However, we failed to find significant correlations between the training-induced improvement of TOEIC scores

and Pre-GM volume (VBM), Pre-FA values (TBSS), or Pre-connectivity of the eight tracts (PDT).

The analyses of the MRI data (Post-1 vs Pre) identified training-induced increases in both GM and WM of the right hemisphere for the TG compared with the CG (significant time-by-group interaction). In the TG compared with the CG, VBM analysis identified significant increases in GM volume only in the right IFGop (5.1 ± 2.3%), corresponding to the neural outcome of plastic changes induced by the L2 vocabulary learning (Fig. 2b). The training-induced increases of GM volume in the right IFGop were correlated with those of TOEIC score ($r = 0.52$, $p = 0.02$).

In the TG, FA values increased from Pre to Post-1 in WM in the right sub-IFGop region (3.6 ± 1.6%; Fig. 2c). This increase in FA showed a strong tendency to correlate with the increase in TOEIC score ($r = 0.33$, $p = 0.05$). Using the PDT analyses, we further discovered learning-related increases in the connectivity parameter of the IFGop-caudate head pathway (4.5% ± 6.4, $F_{(3,82)} = 7.83$, $p = 0.01$) and that of the dorsal pathway (3.7 ± 6.4%, $F_{(3,82)} = 6.72$, $p = 0.01$) in the right hemisphere (2-by-2 repeated-measures ANOVA). The homologous tracts in the left hemisphere did not show the corresponding changes (2.5 ± 4.4%, $F_{(3,82)} = 0.16$, $p = 0.69$ for the left IFGop-caudate connectivity; 2.7 ± 5.2%, $F_{(3,82)} = 0.16$, $p = 0.65$ for the left dorsal pathway connectivity; Fig. 3). There were no significant L2 training-induced connectivity changes in the ventral pathway (2.1 ± 4.0%, $F_{(3,82)} = 0.16$, $p = 0.69$ for the right; 1.9 ± 4.5%, $F_{(3,82)} = 0.26$, $p = 0.70$ for the left) or in the ILF (1.8 ± 5.0%, $F_{(3,82)} = 0.22$, $p = 0.41$ for the right; 2.0 ± 5.1%, $F_{(3,82)} = 0.31$, $p = 0.49$ for the left). The training-induced increases in TOEIC score were correlated with those of the right IFGop-caudate connectivity ($p = 0.02$, $r = 0.59$), but not of the left IFGop-caudate ($p = 0.13$, $r = 0.25$), right dorsal pathway ($p = 0.07$, $r = 0.29$), left dorsal pathway ($p = 0.30$, $r = 0.14$), right ILF ($p = 0.14$, $r = 0.19$), or left ILF connectivity ($p = 0.25$, $r = 0.20$; Fig. 3a,e,f,j,k,o).

We re-examined L2 proficiency and brain structure of the participants in the TG a year after the completion of the training program. One-way repeated-measures ANOVA showed significant differences in GM ($F_{(1.8, 38.6)} = 477.0$, $p < 0.001$), FA ($F_{(1.8, 38.6)} = 233.8$, $p < 0.001$), and right IFGop-caudate connectivity ($F_{(1.8, 38.6)} = 331.0$, $p < 0.001$) across the three time points. Consistent with the aforementioned group comparison, post hoc comparisons (Tukey's HSD test) between Pre and Post-1 demonstrated significant increases in GM volume in the right IFG ($p < 0.001$), FA beneath the right IFGop ($p = 0.03$), and connectivity between the right IFGop and caudate head ($p = 0.02$). Furthermore, the increment of GM volume and FA values from Pre to Post-1 were significantly correlated (Fig. 4a), suggesting parallel plastic changes in GM and WM. Nevertheless, post hoc comparisons (Tukey's HSD test) between Post-1 and Post-2 showed significant decreases in GM volume in the right IFGop (−4.3 ± 2.5%, $p = 0.04$) and FA beneath the right IFG (−3.7 ± 4.2%, $p = 0.04$). In addition, the decreases of GM volume from Post-1 to Post-2 were significantly correlated with those of FA values ($p < 0.001$, $r = 0.65$; Fig. 4b). That means the parallelism of the "negative" changes in GM and WM. We also noticed a non-negligible trend toward decreases in right IFGop- caudate head connectivity (−2.3 ± 8.1%, $p = 0.05$) in the connectivity of the dorsal pathway (−2.0 ± 7.2%, $p = 0.06$). Resultantly, no difference was found between the Pre and Post-2 values for any of the imaging parameters. These findings indicate that in most of the participants, training-induced behavioral values and macroscopic neural structure returned to the pretraining state, suggesting a phenomenon of neural "elasticity." Through the assessment of the