Mental Visual Synthesis is Originated in the Fronto-temporal Network of the Left Hemisphere

Introduction

Mental visual synthesis is the act or power of forming a mental image of something not perceived by the senses or present in reality. It consists of taking parts of our various conceptions and combining them to give new forms and images more selective, more striking, more delightful and more terrible, among other things, than those existing in reality. There is no doubt that our brain generates mental visual synthesis. Therefore, brain networks must be activated and a population of neurons belonging to these brain networks must be activated when a person visualizes synthesized objects in their mind’s eye. Although neuroimaging techniques now have sufficient potential to detect changes in neuronal activity involved in higher cognitive or mental functions, to our knowledge no brain imaging study has been performed that directly investigated the relationship between brain activity and mental visual synthesis. Therefore, in the present study, we used functional magnetic resonance imaging (fMRI) to address the question of which brain areas are involved during mental visual synthesis that is the basis of imagination.

Mental visual imagery is one of the relevant counterparts of mental visual synthesis in the context of human cognitive functions. Mental visual synthesis is the mental capacity for experiencing, constructing, or manipulating mental imagery. Mental visual imagery is defined as the manipulation of visual information arising from memory, giving rise to the experience of seeing through the mind’s eye. In contrast to mental visual synthesis, there are many investigations on mental visual imagery using functional imaging techniques (for reviews, see Mellet et al., 1998; Kosslyn et al., 2001). One of the major controversies in previous functional imaging studies has been whether the primary visual cortex is involved in visual mental imagery. Although many of the previous studies addressed this issue, results of these studies also indicated that not only visual areas but also several brain areas are activated in relation to visual mental imagery and that the brain activation pattern varies according to the object that is mentally visualized. A significant activation of the parieto-occipital areas (often bilaterally) was reported during visual imagery tasks using mental rotation (Mellet et al., 1996), route finding (Roland and Friberg, 1985; Roland et al., 1987; Ghaem et al., 1997) and recall of prelearned three-dimensional scenes (Mellet et al., 2000), patterns (Knauff et al., 2000) and spatial locations (Kawashima et al., 1995). In contrast, temporal-occipital areas were significantly activated when subjects were instructed to imagine construction of concrete objects according to verbal instructions (D’Esposito et al., 1997; Mellet et al., 1998) and based on paired associates (Shallice et al., 1994; Fletcher et al., 1995) and mentally to recall prelearned patterns (Roland and Gulyas, 1995) and letters (Kosslyn et al., 1995). The findings of neuroimaging studies on mental visual imagery are consistent with those of previous neuroimaging studies of visual perception showing functional dichotomy between the dorsal and the ventral pathways according to the spatial and object form perceptions, respectively (for a review, see Ungerleider and Haxby, 1994), which was first demonstrated in monkeys (Miskin et al., 1983). These findings indicate that mental visual imagery and visual perception share common and very specific mechanisms.

In the present study, we aimed to determine brain areas specifically involved in mental visual synthesis. To do this, we compared brain activity measured by fMRI during mental visual synthesis and mental visual imagery conditions. In this study, we assumed that at least the following cognitive processes have to be involved in mental visual imagery: retrieval of knowledge of two different objects from long-term semantic memory based on external cues (words or line drawings); bringing these pieces of information into working memory; holding it on-line; and seeing those objects through the mind’s eye. We also assumed that mental visual synthesis requires one more specific cognitive process in addition to the above-mentioned processes, that is to create a new object by combining and modifying these pieces of information in novel ways. Since we counterbalanced visual inputs and motor outputs between mental visual synthesis and mental visual imagery conditions in this study, a comparison of mental visual
synthesis versus object imagery may indicate brain areas involved in the specific cognitive process for visual synthesis.

**Methods**

**Subjects**

Forty-three young right-handed healthy volunteers, of ages ranging from 18 to 31 years, participated in this study. Their handedness was assessed by the Edinburgh Handedness Inventory. Written informed consent was obtained from each subject in accordance with the guidelines approved by the Tohoku Fukushi University and the Helsinki Declaration of Human Rights, 1975. Brain activity was measured in 23 subjects using fMRI during a word task and the remaining 20 subjects performed a picture task during fMRI measurement.

**Tasks**

The word task consisted of two experimental conditions, that is, object imagery and synthesis conditions and a baseline condition. In the object imagery condition, two different nouns of a common object, e.g. ‘apple’ and ‘table’, were presented at the center of the screen simultaneously for 1 s, followed by a 4 s inter-trial interval. During the inter-trial interval, an eye-fixation point was presented at the center of the screen. The subjects were instructed to imagine the appearance of the real objects in turn during the inter-trial interval and were asked to press a button using the right hand when they succeed to imagine two objects. In the synthesis condition (Fig. 1A), a compound word, which consisted of two nouns but did not indicate a real object, e.g. ‘television–rock’, was presented for 1 s at the center of the screen, followed by a 4 s inter-trial interval. The subjects were instructed to imagine a novel object during the inter-trial interval and were asked to press a button when they succeed in visualizing an object. Actual verbal instruction for synthesis condition was as follows; ‘we ask you to create a vivid image of an object, which does not exist in the real world, in your mind from an artificial word consisting of two nouns of familiar objects’. In the base line condition, the subjects were asked to gaze at an eye fixation point. The visual inputs, that is, the presented words, were counterbalanced between the object imagery and synthesis conditions.

The picture task also consisted of two experimental conditions, that is, object imagery and synthesis conditions and a baseline condition. In the object imagery and synthesis conditions, two different achromatic line drawings of common objects were presented for 1 s at the center of the screen, followed by a 4 s inter-trial interval. During the inter-trial interval, an eye-fixation point was presented at the center of the screen. In the object imagery condition, the subjects were asked to imagine the appearance of real object in turn during the inter-trial interval and were asked to press a button when they succeed in visualizing two objects. In the synthesis condition (Fig. 1B), the subjects were instructed to create a novel object in their minds by compounding the two line drawings, to imagine this object during the inter-trial interval and to press a button when they succeed in visualizing this object. Actual verbal instruction for the synthesis condition was the same as for the word task. The presented line drawings were generated from the same words used in the word task. The baseline condition was the same in the word task. The visual inputs and motor outputs were counterbalanced between the object imagery and synthesis conditions.

Prior to the fMRI scan, subjects were trained with both object imagery and imagination conditions. The reaction time, that is, duration between the start of the inter-trial interval and pressing the button, was measured during fMRI measurements using a personal computer. A debriefing of the subject was held immediately after each fMRI scan.

**fMRI Measurements**

$T_1$-weighted structural images were acquired for each volunteer, using a 1.5 T Siemens Vision plus scanner ($T_R = 9.7$ ms, echo time = 4 ms, FA = 12°, FOV = 250 mm, pixel size = $1.25 \times 0.98$ mm, matrix = $175 \times 256$; Siemens Magnetron Vision, Erlangen, Germany). To measure neural responses, gradient echo, echo-planar $T_2^*$-weighted images with blood oxygenation level-dependent (BOLD) contrast were acquired by GE-EPI ($T_E = 60$ ms, FA = 90°, matrix = $64 \times 64$, FOV = $256 \times 256$ mm, slice thickness 3 mm, gap = 2.25 mm) covering the entire brain at a $T_R$ of 5 s. Forty-two axial slices were oriented over the whole brain. To eliminate $T_1$ equilibration effects, four preliminary scans were acquired and the others subsequently discarded. Thereafter, scans were acquired continuously every 5.0 s.

In the present study, a conventional block design consisting of ten alternations of a 60 s ON block and a 20 s baseline condition block was used for each task. Each of the two experimental conditions was assigned to the ON block five times in the same order for each subject (e.g. ABCABCABABCA) and the order of the experimental conditions was counterbalanced across the subjects. Therefore, each ON block consisted of 10 trials for each task.

**Image Data Processing and Statistical Analyses**

All preprocessing procedures and data analyses were performed using Statistical Parametric Mapping 99 (SPM99) software (Wellcome Department of Cognitive Neurology, London, UK) implemented on MATLAB (Mathworks Inc., Natick, MA). First, slice time adjustment was performed to correct for differences in acquisition time among 34 slices of each scan. Secondly, to correct for artifacts caused by small head movements, images from each subject were realigned to the first image and resliced using a sinc interpolation. Thirdly, a mean image created from the realigned volumes was coregistered with the structural $T_1$ volume and the structural volumes were spatially normalized to a standard template in the space of Talairach and Tournoux (1988) using nonlinear basis functions. Fourthly, the derived spatial transformation was applied to the realigned $T_2^*$ volumes, which were then spatially smoothed with a three-dimensional isotropic Gaussian Kernel (12 mm full width half-maximum). This improves the signal-to-noise ratio and accommodates for residual variations in functional neuroanatomy that usually persist between subjects after spatial normalization.

Condition effects at each voxel were estimated according to this general linear model and regionally specific effects were compared using linear contrasts. Global changes were adjusted by a proportional scaling and low-frequency confounding effects were removed using an appropriate high-pass filter. Voxel values for each contrast yielded a statistical parametric map of the $t$-statistic (SPM $t$) and were subsequently transformed to the unit normal distribution (SPM $Z$). Then, inter-subject maps were produced by performing one-sample $t$-tests to identify voxels that survived the statistical threshold of $P < 0.05$ after correction for multiple comparisons. Finally, the resulting activation maps were constructed and superimposed onto stereotactically standardized $T_1$-weighted MR images.

A set of ROIs was anatomically defined for each subject on the regions that were significantly activated in the synthesis versus object

![Figure 1](https://example.com)
imagery comparisons. For each subject and each ROI, a mean time series, averaged across activated voxels in the region and across all runs, was calculated. We then compared the mean amplitude of fMRI signal for each condition.

Results

Task Performance
Immediately after the fMRI measurements, the subjects were asked to sketch the appearance of objects they imagined during the synthesis conditions. Examples of these sketches are shown in Figure 1.

The mean (SEM) reaction times during object imagery and synthesis conditions of the word task were 1.3 s (0.1) and 1.4 s (0.1), respectively, with no statistically significant differences (paired t-test). Those during object imagery and synthesis conditions of the picture task were 2.5 s (0.2) and 2.7 s (0.2), respectively, with no statistically significant differences (paired t-test). The reaction time in the picture task was statistically significantly longer than that in the word task (ANOVA, \( P < 0.001 \)).

The mean (SEM) percentages of successful trials of the word task were 98.0% (2.8) and 95.0% (2.7) during object imagery and synthesis conditions with no statistically significant differences (paired t-test). Those during object imagery and synthesis conditions of the picture task were 93.4% (4.0) and 90.0% (5.0), respectively, with no statistically significant differences (paired t-test). The mean percentage of successful trials was not statistically significant between the word and picture tasks (ANOVA).

Brain Activation
Table 1 and Figure 2 summarize brain areas activated during the word task. The object imagery versus baseline comparison revealed the significant activation of the inferior frontal and inferior temporal cortices of the left hemisphere, anterior cingulate cortex, as well as bilateral cerebellum. The synthesis versus baseline comparison showed the significant activation

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Talairach coordinates and ( T )-value of peak activation during the word task</td>
</tr>
<tr>
<td>Anatomical name</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Object imagery versus baseline</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior temporal</td>
</tr>
<tr>
<td>Lt cerebellum ant lobule</td>
</tr>
<tr>
<td>Rt anterior cingulate</td>
</tr>
<tr>
<td>Rt cerebellum anterior lobule</td>
</tr>
<tr>
<td>Rt cerebellum posterior lobule</td>
</tr>
<tr>
<td>Rt cerebellum posterior lobule</td>
</tr>
<tr>
<td>Synthesis versus baseline</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior temporal</td>
</tr>
<tr>
<td>Lt intraparietal</td>
</tr>
<tr>
<td>Lt parahipocampal</td>
</tr>
<tr>
<td>Rt parahipocampal</td>
</tr>
<tr>
<td>Rt striatum</td>
</tr>
<tr>
<td>Anterior cingulate</td>
</tr>
<tr>
<td>Rt cerebellum anterior lobule</td>
</tr>
<tr>
<td>Synthesis versus object imagery</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior frontal</td>
</tr>
<tr>
<td>Lt inferior temporal</td>
</tr>
<tr>
<td>Rt striatum</td>
</tr>
<tr>
<td>Lt striatum</td>
</tr>
<tr>
<td>Rt cerebellum posterior lobule</td>
</tr>
</tbody>
</table>

Stereotaxic coordinates (in millimeters) identify the location of the maxima of haemodynamic responses corresponding to the atlas of Talairach and Tournoux (1988). Lt and Rt indicate left and right hemisphere, respectively. Volumes are in mm\(^3\).
of the inferior frontal, inferior temporal and intraparietal cortices of the left hemisphere, anterior cingulate cortex, bilateral parahippocampal cortices, right insula and the right cerebellum. The synthesis condition more significantly activated the left inferior prefrontal cortex, left inferior temporal cortex, bilateral striatum and the right cerebellum than the object imagery condition.

Table 2 and Figure 2 summarize brain areas activated during the picture task. The object imagery versus baseline comparison showed significant activation of the left inferior frontal cortex, anterior cingulate cortex and bilateral cerebellum. Although subjects looked at the line drawing pictures during the object imagery condition, activation of the primary visual cortex did not reach a statistically significant level. The synthesis versus baseline comparison revealed the significant activation of the bilateral inferior frontal, left inferior temporal, left intraparietal, left fusiform and anterior cingulate cortices, as well as the right striatum and right cerebellum. The synthesis condition more significantly activated the left inferior frontal, bilateral inferior temporal, bilateral intraparietal and left lateral occipital cortices than the object imagery condition. A larger extent of activation was evoked by the picture task compared with the word task in each comparison.

It should be noted that the same two areas in the left inferior frontal cortex, one in the anterior region, the other in the posterior region and an area in the posterior part of the left inferior temporal cortex were significantly activated during the synthesis condition compared with the object imagery condition during both tasks (Fig. 2). The anterior and posterior regions of the inferior frontal gyrus are probably located in the orbital and opercular parts of the inferior frontal gyrus, respectively. The magnitude of activation in these areas during each condition was not statistically significant (Student’s t-test) between the picture and word tasks (Fig. 3).

Discussion
Our results indicate that the neuronal network consisting of the inferior frontal and the inferior temporal cortices of the left
hemisphere plays an important role in creating an entirely new image in one’s mind.

In this study, we assumed that in order to perform the synthesis condition of each task, at least following the three cognitive processes are involved: (i) retrieval of information, that is, knowledge of two different objects, from long term semantic memory provided by external cues (words or line drawings) and bringing this information into working memory and holding it on-line; (ii) creation of a new object by combining and modifying this information in novel ways; and (iii) visualize that object through the mind’s eye. We also assumed that processes (i) and (iii) are necessary to perform the object imagery condition as well. Therefore, visual inputs and motor outputs were counterbalanced between the object imagery and synthesis conditions during each task; a comparison of synthesis versus object imagery may indicate brain areas involved in the creation of entirely new objects by combining two different representations of the objects.

**Hemispheric Asymmetry**

We found that both object imagery and synthesis evoked the left-lateralized activation of the fronto-temporal network. The similar hemispheric asymmetry was suggested by brain imaging studies of visual imagery of familiar objects (D’Esposito et al., 1997; Ishai et al., 2000, 2002). On the other hand, the activation of the right inferior temporal cortex has also been reported during the mental imagery of complex objects or schematic objects (Mellet et al., 1996) and letters (Kosslyn et al., 1993). Mellet et al. (1998) reviewed previous brain imaging studies and concluded that functional lateralization during mental imagery can depend on two different characteristics of mental images to be generated: complexity, which would modulate the degree of involvement of the right hemisphere; and lexicality, which would drive the left hemisphere participation.

Farah (1986) found that left hemisphere was better at using imagery to prime perceptual recognition by a psychophysical study of normal subjects, although Kosslyn et al. (1985) found that in split-brain patients this left-hemisphere advantage was selective, occurring only for multipart images and that both hemispheres were equally adept at generating single-unit images of an overall shape. Kosslyn et al. (1995) suggested that left hemisphere more effectively generates images by arranging parts according to descriptions (using categorical spatial relat-
The previous neuroimaging investigations suggested that there is no doubt the inferior frontal cortex is involved in semantic processing. One of the functional roles of the inferior frontal cortex derived from the results of previous neuroimaging studies is semantic generation. Previous positron emission tomography (PET) studies of verb generation tasks, that is, to generate semantically appropriate verbs from the presented nouns, reported activation of the left inferior frontal cortex (Petersen et al., 1988; Raichle et al., 1994; Tatsumi et al., 1999). This area is also activated when subjects name colors (Wiggs et al., 1999) or generate action words (Martin et al., 1995) from visually presented line drawings of objects. Vandenberghe et al. (1996) also suggested that the left inferior frontal cortex is commonly activated when subjects make semantically based decisions on both presented words and pictures. Another functional role of the inferior frontal cortex determined from brain imaging studies is retrieval of semantic knowledge. Previous PET studies using picture naming paradigms indicate involvement of the left inferior frontal cortex in category-specific retrieval (Damasio et al., 1996; Martin et al., 1996; Grabowski et al., 1997). fMRI studies of semantic category generation and semantic memory tasks also activated this area (Shaywitz et al., 1995; Gabrieli et al., 1998; Cossom et al., 1999). In addition, some studies indicated that the inferior frontal cortex is involved in not only for retrieval but also semantic encoding (Demb et al., 1995; Owen et al., 1996; Nyberg et al., 1996). Demb et al. (1995) suggested that the inferior frontal cortex is part of a semantic executive system that contributes to the on-line retrieval of semantic information, similar to executive systems in a model or working memory (Baddeley, 1986; Goldman-Rakic, 1987). However, the inferior frontal cortex was not activated during the Tower of London task, which required motor planning and working memory (Baker et al., 1996), or during non-verbal encoding tasks (Golby et al., 2001). The results of previous studies suggest that the inferior frontal cortex is involved in semantic encoding and retrieval. The other function is selection of semantic knowledge. The prefrontal cortex has been argued to produce flexible and context-sensitive responses (Cohen and Servan-Schreiber, 1992) and to mediate the selection of action by the weighing of information active in working memory (Kimberg and Farah, 1993). From the results of an fMRI study, Thompson-Schill et al. (1997) found that demands on the inferior frontal cortex are high in any task that requires selection among competing sources of information in working memory to guide a response. Since semantic generation, retrieval of semantic knowledge from presented words or pictures and selection of semantic knowledge are important cognitive processes for mental visual synthesis, we assumed that the results of our study are in agreement with those of previous imaging studies and that several different semantic modules in the inferior frontal cortex are responsible to mental visual synthesis.

Although, in this study, we were unable to determine functional dissociations of the anterior and posterior regions of the left inferior frontal cortex, some previous neuroimaging studies indicated interesting functional differences between the two regions. Poldrack et al. (1999) and Heim et al. (2003) suggested from the results of their fMRI studies that the anterior region subserves semantic processing and the posterior region subserves phonological processing. Bokde et al. (2001) obtained consistent results from functional connectivity ana-

**Figure 3.** Mean amplitude of fMRI signal in significantly activated areas in synthesis versus object imagery conditions, respectively. Error bars indicate standard error of the mean (SEM). MFC, middle frontal cortex; IFC, inferior frontal cortex; IPC, intraparietal cortex; ITC, inferior temporal cortex; Str, striatum; CBp, posterior lobe of the cerebellum; L, left hemisphere; R, right hemisphere.
lysis using fMRI. Fletcher and Henson (2001) summarized from the literature review that the anterior region is involved in the semantic contribution to successful encoding and the posterior region is involved in selection from various semantic attributes. At the moment, we cannot explain for the differences in functional organization of these two regions, although we can argue that the two regions may have different functional roles and that the both areas are involved in mental visual synthesis.

In summary, since inferior frontal cortex was activated during the object imagery as well as synthesis conditions and was more significantly activated during the synthesis condition, the difference in the inferior frontal activation between these two conditions may well be resulted from the difference in the semantic demands of the tasks.

**Inferior Temporal Cortex**

In this study, the posterior part of the inferior temporal cortex was more activated during the synthesis condition than during the imagery condition in both tasks. Activation of the same area was reported in the previous brain imaging studies during retrieval of object features (Haxby et al., 1994; Kohler et al., 1995; Moscovitch et al., 1995; Martin et al., 1996; Owen et al., 1996) and encoding of object features (Owen et al., 1996). This area has been considered as a target of the ventral stream (Miskin et al., 1983) originating from the visual cortex, which is specialized for processing object forms (for reviews, see Goodale, 1993; Ungerleider and Haxby, 1994). Previous lesion studies of humans revealed that the posterior temporoparietal damage impairs mental-image generation (Sergent, 1996; Tippet, 1992; Trojano and Grossi, 1994; Farah, 1995). Therefore, it is argued that the inferior temporal cortex is involved in object mental imagery and perception (Mellet et al., 1998). The involvement of the left inferior temporal cortex in semantic processing was also reported in previous brain imaging studies (Tranel et al., 1997; Martin et al., 2000). Since one can easily build a mental image based on a verbal description (Denis and Cocude, 1992), or provide an oral description of a mental image, we may argue that strong interactions exist between language and mental imagery that permit an easy transfer from one representation mode to the other and that the interactions may take place in the inferior temporal cortex.

It should be noted that a recent fMRI study by Bokde et al. (2001) indicated strong functional connectivity between anterior or posterior regions of the left inferior frontal cortex and the posterior part of the inferior temporal cortex during phonological processing of words. Their findings support our results, since we noted that the inferior temporal cortex was co-activated with the left inferior frontal cortex during the synthesis condition.

**Notes**

This study has been supported by JST/RISTEX, R&D promotion scheme for regional proposals promoted by TAO, a Grant-Aid for Scientific Research on Priority areas (C) – Advanced Brain Science Project – from MEXT and the 21st Century Center of Excellence (COE) Program (Ministry of Education, Culture, Sports, Science and Technology) entitled ‘A Strategic Research and Education Center for an Integrated Approach to Language and Cognition’ (Tohoku University).

Address correspondence to Ryuta Kawashima, NIChe, Tohoku University, Aramaki aza Aoba, Aobaku, Sendai 980-8579, Japan. Email: ryuta@idac.tohoku.ac.jp.

---

**References**


