Learning second language vocabulary: Neural dissociation of situation-based learning and text-based learning

Hyeonjeong Jeonga,⁎, Motoaki Sugiarua, Yuko Sassa, Keisuke Wakusawab, Kaoru Horiec, Shigeru Satoc, Ryuta Kawashimaa

aDepartment of Functional Imaging, Institute of Development, Aging and Cancer, Tohoku University, Seiryo-machi 4-1, Aoba-ku, Sendai 980-8575, Japan
bDepartment of Pediatrics, Tohoku University Graduate School of Medicine, Sendai, Japan
cGraduate School of Intercultural Studies, Tohoku University, Sendai, Japan

Abstract

Article history:
Received 18 September 2009
Revised 18 November 2009
Accepted 8 December 2009
Available online 16 December 2009

Keywords:
Word learning
Second language
Encoding
Retrieval
Social context
Text
fMRI

Second language (L2) acquisition necessitates learning and retrieving new words in different modes. In this study, we attempted to investigate the cortical representation of an L2 vocabulary acquired in different learning modes and in cross-modal transfer between learning and retrieval. Healthy participants learned new L2 words either by written translations (text-based learning) or in real-life situations (situation-based learning). Brain activity was then measured during subsequent retrieval of these words. The right supramarginal gyrus and left middle frontal gyrus were involved in situation-based learning and text-based learning, respectively, whereas the left inferior frontal gyrus was activated when learners used L2 knowledge in a mode different from the learning mode. Our findings indicate that the brain regions that mediate L2 memory differ according to how L2 words are learned and used.

© 2009 Elsevier Inc. All rights reserved.

Introduction

The manner in which linguistic knowledge, such as words and grammar, is learned and used is an important aspect of second language (L2) acquisition. Learning (encoding) and use (retrieval) are highly interdependent, as the cognitive operations needed for learning a word have a profound impact on subsequent retrieval (Morris et al., 1977; Tulving and Thomson, 1973). According to the cognitive neuroscience literature, the processing demands of encoding various items, such as words, faces, and pictures, critically determines the brain regions that are involved in the successful retrieval of these items (Johnson and Rugg, 2007; Nyberg et al., 2000; Otten et al., 2002). However, no studies have investigated how the mode of L2 learning influences which brain regions mediate retrieval.

In the present study, we used functional magnetic resonance imaging (fMRI) to investigate the neural mechanisms underlying encoding and retrieval of L2 word knowledge. We first examined how different types of learning L2 vocabulary affected cortical processing during retrieval. We focused on text-based learning, defined as encoding new words accompanied by written translations, and situation-based learning, defined as encoding words in real-life situations that provided a functional meaning for the words. We hypothesized that dissociation in neural representation exists between situation-based and text-based learning. In the situation-based learning condition, learners have to extract the meaning of new words in many different contexts by observing and integrating multiple signals such as the actions and intentions of the speaker using them. This type of learning appears to be ubiquitous in infant first language (L1) acquisition, which requires social cognitive skills (Tomasetto and Akhtar, 1995; Yu et al., 2005). In the text-based learning condition, learners largely rely on rote associative memorization between new words and their translations in L1. This is known as word-paired associate learning, a typical way of L2 vocabulary learning (Nation, 2001). These two learning types may require distinct cognitive operations and thus may produce qualitatively different neural processing during subsequent use.

Additionally, we were interested in how learners apply L2 word knowledge that has been learned in a particular mode (text-based) to another mode (situation-based). The flexible application of L2 vocabulary knowledge into various contexts, such as real-life communication, multiple-choice tests, and translation is critical for successful L2 acquisition. We hypothesized that controlled cognitive processing was required when learners retrieve a word in a mode different from the one in which it was encoded. We further hypothesized that the left prefrontal cortex was involved in this process because it mediates cognitive control of semantic memory.

⁎ Corresponding author. Fax: +81 22 717 7988.
E-mail address: jeong@idac.tohoku.ac.jp (H. Jeong).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
(Badre et al., 2005; Wagner et al., 2001), strategic access to knowledge (Anderson et al., 2004; Sohn et al., 2003), and general cognitive control for task switching (Miller and Cohen, 2001). Previous neuroimaging studies have observed activation of the left prefrontal cortex during various cognitive control tasks, but it is not clear how newly acquired L2 knowledge is processed when the encoding and retrieval conditions are different.

In the present experiment, we recruited Japanese university students with no prior knowledge of Korean, the target language. Subjects were asked to learn the meanings of spoken Korean words in either a situation-based (SL) or text-based (TL) learning condition. We prepared short movie clips for both conditions: in TL, the words were spoken by a person holding a board on which the Japanese translation was written; in SL, the words were spoken in a social situation. Similar movie clips were presented to the subjects in the retrieval task, and the spoken L2 (i.e., Korean) words were either correct or incorrect translations of the shown L1 (i.e., Japanese) words or usages of the words in the situation. The subjects were asked to judge whether the spoken L2 word was correct or incorrect while their brain activity was scanned using fMRI.

The TL and SL words were both tested under situation-based (S) and a text-based (T) conditions. The experiment had a two-factorial design with the following four conditions: SL, SSL, SLT, and ST. The effect of encoding was investigated by comparing words encoded under the situation (SS) and text (ST) conditions. The interaction between encoding and retrieval was investigated by comparing the mismatched conditions (SSL and SLT), which required different cognitive operations between encoding and retrieval, with the matched conditions (SS and ST). The L1 control task was also included in the situation-based and text-based retrieval conditions (SLS and SlT).

The present study used the framework of the general cognitive model of memory, a novel approach for neuroimaging studies of L2 acquisition (Morris et al., 1977; Tulving and Thomson, 1973). Although previous neuroimaging studies of L2 have investigated the neural correlates of novel vocabulary acquisition (Breitenstein et al., 2005; Mestres-Misse et al., 2008; Raboyeau et al., 2004), this is the first study to compare learning types and to examine the relationship between encoding and retrieval.

Methods

Subjects

The subjects were 44 healthy right-handed native Japanese speakers. According to the behavior and imaging data criteria (reported below), we tested 31 subjects (21 men and 10 women) with an average age of 21.6 (range = 18–26, SD = 2.31). The subjects were undergraduate and graduate students at Tohoku University in Sendai, Japan. They had normal language development and experienced no difficulties in learning language. None had prior experiences in learning Korean. They had normal vision, and none had a history of neurological or psychiatric illness. Handedness was evaluated using the Edinburgh Handedness Inventory (Oldfield, 1971). Written consent was obtained for each participant, and the study met all criteria for approval by the Ethical Committee of Tohoku University.

Creation of the stimuli

We made movie clips for the SL, TL, and retrieval test conditions. In the movie clips, Korean words were spoken in actual life situations (Fig. 1a) or were spoken by a person showing a board on which the Japanese (L1) translation of the word was written (Fig. 1b). The movies were created using the following procedure: we selected 48 Korean words including verbs, adjectives, and greetings frequently used in daily life. Then, eight different scenarios were developed for each word. In the situation-based learning movies, seven native Korean speakers or Japanese speakers fluent in Korean participated, and one or two of them acted and used the target word in each situation. For example, to express the target word “Dowajo,” meaning help me in English, an actor tried to move a heavy bag and asked the other actor for help. To create as natural a situation as possible, we made the movies in various environments such as the school, office, home, park, station, train, restaurant, and parking lot. Each movie lasted 5 s. We pre-tested 15 volunteers, who were similar to the study subjects and who had never studied Korean, to select the movies that best expressed the meaning of the target word. The volunteers were asked to watch eight movie clips, guess the meanings of the target words, and write them down in Japanese. They were also asked to point out ambiguous movies in which the meaning was difficult to guess. Following this procedure, we selected 24 words for which the meaning was correctly guessed by all volunteers and seven situation-movie clips for each word. Korean native speakers who acted in the situation movies also participated in the text movie. Under this condition, each of seven native Korean speakers spoke a target word while holding a board on which the L1 translation was written. Seven text-movie clips were made for each word. Five of seven movie clips were used in the learning session, and the resting two were used for the test session. The latter clips were truncated to 3 s in the length, which sufficiently covers the part of the film where the word is spoken and necessary contextual information is provided.

Learning session

The subjects participated in both the learning and testing sessions. The 24 target words were divided into two groups to ensure that the number of verbs, adjectives, and greeting expressions were equal. The target words were used for either the SL or TL condition and were
count-balanced across subjects. In the learning session, the subjects were asked to watch and memorize the 12 words presented under each learning condition, and the session continued until all 24 target words were memorized. We used five movie clips for each target word for both the TL and SL conditions. All the movie clips for both the learning conditions were bundled in a single sequence of movie clips. The subjects were given headphones, and they viewed the sequence of movie clips on a computer. Each subject repeatedly viewed the entire sequence and was not allowed to see only a part of the sequence. When the target word was first presented, the pronunciation was written on the bottom of each movie clip to help subjects learn the correct pronunciation. The learning session took about 3 h. To avoid the effect of rote rehearsal memory, we asked subjects to perform an off-task, which was another fMRI experiment unrelated to this experiment (Wakusawa et al., 2007), for about half an hour. They were given a half-hour rest following the off-task.

Testing session (fMRI task)

During the test session, the subjects were asked to judge whether or not the spoken words were a correct translation of the shown L1 words or correct usages of the words in the given situations in the movies. Target words that were learned in the SL and TL conditions were tested both in the S and the T conditions. In the S test, Korean native speakers acted out correct or incorrect scenarios using the learned Korean words. In the T test, a Korean native speaker spoke the word while holding a board with a correct or incorrect Japanese (L1) translation of the word. The study consisted of the following four test conditions: (1) \( S_L S \): SL words were tested in a situation-based test; (2) \( S_L T \): SL words were tested in a text-based test; (3) \( T_L S \): TL words were tested in a situation-based test; and (4) \( T_L T \): TL words were tested in a text-based test (Fig. 2). The \( S_L S \) and \( T_L T \) conditions were matched between the learning and test phase, and the \( S_L T \) and \( T_L S \) conditions were mismatched. Movie clips similar to those for the L2 conditions were prepared for both the text-based and situation-based tests in the subjects’ first language (\( L_1 S \) and \( L_1 T \)). This allowed comparison of the activation pattern between L2 and L1 word retrieval. The different verbs, adjectives, and greeting words were used in the control task. Both the situation-based tests (\( S_L S \), \( T_L S \), and \( L_1 S \)) and text-based tests (\( S_L T \), \( T_L T \), and \( L_1 T \)) consisted of 24 trials in which target words were correctly used and 12 trials in which target words were incorrectly used. All movie clips lasted 3 s. The L2 and L1 task conditions were presented separately by blocks. At the beginning of each block, the subjects were presented with a visual signal lasting 5 s indicating whether the next task was L1 or L2. Every stimulus in each block was pseudo-randomly presented by the event-related experimental paradigm with a 1.5- to 5-s interval between movie clips. The order of blocks was counter-balanced across subjects. The entire experiment took 1248 s. The subjects were required to respond with an optical button held between their right index and middle fingers as soon as they knew the answer. Each response was digitally recorded.

Data acquisition

Forty-four gradient-echo images (echo time = 50 ms, flip angle = 90°, slice thickness = 2.2 mm, slice gap = 0.77 mm, field of view = 192 mm, matrix size = 64 x 64) covering the entire cerebrum were acquired at a repetition time of 4000 ms using an echo planar sequence and a magnetic resonance scanner (1.5 T Siemens Symphony, Erlangen, Germany). Excluding three dummy scans for stabilization of the T1 saturation effect, 309 volumes were acquired for each subject. Additionally, a T1-weighted data set was acquired.

---

**Fig. 2.** Examples of the retrieval test for each learning condition. A total of six experimental conditions (\( S_L S \): situation-based test of words learned from a situation; \( S_L T \): text-based test of words learned from a situation; \( T_L S \): situation-based test of words learned from text; \( T_L T \): text-based test of words learned from text; \( L_1 S \): situation-based test of L1 words; \( L_1 T \): text-based test of L1 words).
The following preprocessing procedures were performed using Statistical Parametric Mapping (SPM5) software (Wellcome Department of Imaging Neuroscience, London, UK) and MATLAB (MathWorks, Natick, MA, USA): adjustment of acquisition-timing across slices, correction for head motion, coregistration to the anatomical image, spatial normalization using the anatomical image and the MNI template, and smoothing using a Gaussian kernel with a full-width at half-maximum of 9 mm. Imaging data that showed more than 3 mm excessive motion within the scanner were excluded from the statistical analysis.

**fMRI analysis**

A conventional two-level approach for event-related fMRI data was adopted using SPM5. A voxel-by-voxel multiple regression analysis of expected signal changes for each condition, which were constructed using the hemodynamic response function provided by SPM5, was applied to the preprocessed images for each subject. Correct trials (SLS, SLT, τ1S, τ1T, and SLS), incorrect trials, and unsuccessful trials in which the subject gave the wrong answer under all conditions were modeled separately. This was accomplished by excluding the brain activation trials in which a subject could not retrieve the correct word. Statistical inference on contrasts of parameter estimates was then performed using a second-level between-subject (random effects) model using one-sample t-tests.

In the second-level analysis, we compared brain activation in correct trials between the conditions. The statistical threshold in the voxelwise analysis, assuming a search area of the entire brain, was set at corrected \( P < 0.05 \) for family-wise error (FWE). A masking procedure was applied to limit the analysis of the main contrast to the areas showing significant activation in the mask contrast at a liberal statistical threshold \( P < 0.05 \) for height, without correction for multiple comparisons.

First, we analyzed the L2 tasks \((\tau_1S, \tau_1T, \tau_2S \text{ and } \tau_2T)\) compared with the control (L1) task \((\tau_1S \text{ and } \tau_1T)\). This comparison was conducted within the voxels identified by the mask contrasts \(\tau_2S - \tau_1S, \tau_2T - \tau_1T, \tau_2S - \tau_1S, \text{ and } \tau_2T - \tau_1T\). The activation depicted by this comparison reflected the neural activation underlying retrieval of L2 vocabulary relative to that of L1. Then, to identify the main effect of learning types, brain activation was compared between words in the situation-based learning condition \((\tau_2S \text{ and } \tau_2T; \text{ SL})\) and those in the text-based learning condition \((\tau_1S, \tau_1T; \text{ TL})\) during retrieval of words to be learned. Direct comparisons of brain activation between SL and TL conditions were masked by the contrasts \(\tau_2S - \tau_1S \text{ and } \tau_2T - \tau_1T \text{ for the SL-} \text{ TL comparison and the contrasts } \tau_1S - \tau_2S \text{ and } \tau_1T - \tau_2T \text{ for the TL-} \text{ SL comparison. \text{ The activation depicted by these comparisons reflected the neural activation mediating the retrieval of target words encoded differently in the situation and text conditions. Then, to examine the interaction between learning types and testing types, the mismatched condition } \tau_1S + \tau_2T \text{ was compared with the matched condition } \tau_2S + \tau_1T\). This comparison was masked by the contrasts \(\tau_1S - \tau_2S \text{ and } \tau_1T - \tau_2T\). The brain activation observed in this comparison reflects the mechanisms underlying the retrieval and application of the words in different cognitive stages from the words encoded. The resulting activation maps were superimposed on the standard T1-weighted MR image.

Finally, we performed region of interest (ROI) analyses in brain areas obtained from the abovementioned statistical analyses. We extracted the parametric estimates of the activation peak of observed brain areas under each condition and tested significance at a region-level threshold \( P < 0.05 \), without correction for multiple comparisons. A two-way repeated-measure ANOVA was conducted to evaluate the effect of \(a\) the type of learning (SL, TL, and L1), \(b\) the type of test (S and T), and the interaction effect between \(a\) and \(b\).

### Results

#### Behavioral data

The correct responses and reaction times (RT) were recorded. Subjects who scored below 60% on one or more of the four L2 conditions were excluded. Those who responded at 75% of the average accuracy rate under the four L2 conditions were also excluded from the analysis. Thus, data from 31 subjects were analyzed. Table 1 shows the mean percentage of correct responses and RT for all conditions. A two-way repeated-measure ANOVA was conducted to evaluate the effect of the type of learning, type of test, and their interaction on accuracy and RT. The main effect of learning type and the learning × test interaction was significant for both accuracy \([F(1,65, 49.68) = 40.14, P < 0.001 \text{ and } F(2,60) = 10.35, P < 0.001\), respectively] and RT \([F(1.67, 50.23) = 62.62, P < 0.001 \text{ and } F(2, 60) = 18.22, P < 0.001\), respectively].

Paired-samples t-tests were conducted to follow up the significant interactions. The response under the \(\tau_2S\) condition was significantly less accurate than that under the \(\tau_2S\) condition \([t(30) = 4.88, P < 0.001]\), but no difference was found between the \(\tau_2S\) and \(\tau_1T\) conditions. The reaction time was significantly longer under the matched conditions \([t(30) = 4.99, P < 0.001]; \tau_1T \text{ [t(30) = 2.88, } P = 0.007]\).

#### Imaging data

The brain regions involved in L2 vocabulary retrieval were identified by directly comparing the L2 tasks \((\tau_2S, \tau_2T, \tau_1S \text{ and } \tau_1T)\) and the control L1 tasks \((\tau_1S \text{ and } \tau_1T)\). The L2 vocabulary retrieval activated the left inferior, middle, and superior frontal areas, anterior cingulate cortex, temporal areas, parietal lobule, bilateral insula, and basal ganglia (Table 2).

The main effect of the type of learning on the retrieval task was found by comparing brain activation during retrieval processing under both learning conditions \((\text{SL: } \tau_2S, \tau_2T \text{ and TL: } \tau_1S, \tau_1T)\). Retrieval of SL words elicited significantly more activity in the right supramarginal gyrus than the TL words (Fig. 3a and Table 3). Examination of the activation profile of the right supramarginal gyrus revealed a significant main effect of learning type \([x, y, z: 68, −20, 18], F(1,26, 37.86) = 7.46, P = 0.006]\). Pairwise comparisons following the significant main effect of learning were conducted using the Bonferroni correction. SL and L1 words elicited significantly more activity in the right supramarginal gyrus than did TL words \(P < 0.001, P = 0.024\).

In contrast, retrieval of TL words \((\tau_1S \text{ and } \tau_1T)\) elicited greater activation of the left middle frontal gyrus compared to SL words \((\tau_2S \text{ and } \tau_2T; \text{ Fig. 3b and Table 3})\). Examination of the activation profile in this area revealed a significant main effect of learning type \([x, y, z: −32, 48, 8], F(1.38, 41.56) = 6.43, P = 0.009]\). TL words elicited significantly more activity in the left middle frontal gyrus than did SL and L1 words \(P < 0.001, P = 0.028\) in a post hoc analysis using the Bonferroni correction.

Finally, the brain areas involved in an interaction between encoding (i.e., type of learning) and retrieval (i.e., test) were

### Table 1

Mean accuracy and reaction time.

<table>
<thead>
<tr>
<th></th>
<th>Test type</th>
<th>SL words</th>
<th>TL words</th>
<th>L1 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>S</td>
<td>83.01 (7.36)</td>
<td>76.77 (7.11)</td>
<td>88.25 (4.26)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>88.23 (8.25)</td>
<td>91.52 (6.18)</td>
<td>98.08 (4.17)</td>
</tr>
<tr>
<td>Reaction time</td>
<td>S</td>
<td>3.08 (0.36)</td>
<td>3.27 (0.32)</td>
<td>2.79 (0.27)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>2.66 (0.35)</td>
<td>2.54 (0.31)</td>
<td>2.23 (0.32)</td>
</tr>
</tbody>
</table>

The mean accuracy (percentage of correct responses) and reaction time (seconds) are shown. Standard deviations (SDs) are shown in parentheses.
identified by comparing the mismatched conditions (\(S + T\) and \(S\)) and matched conditions (\(S + T\) and \(S\)). The mismatched conditions elicited significantly more activity in the left triangular part of the left inferior frontal gyrus (IFG) (Fig. 4 and Table 4). Examination of the activation profile in the two peaks of the left triangular part of the IFG revealed a significant learning \(\times\) test interaction \([x, y, z; −50, 14, 28], F(2, 60) = 14.63, P < 0.001; [x, y, z; −50, 14, 28], F(2, 60) = 13.40, P < 0.001\).

**Table 2**

Brain areas associated with retrieving L2 words.

<table>
<thead>
<tr>
<th>Structure</th>
<th>(x, y, z)</th>
<th>T-value</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left opercular part of inferior frontal gyrus</td>
<td>−46, 8, 28</td>
<td>7.14</td>
<td>310^a</td>
</tr>
<tr>
<td>Left precentral sulcus</td>
<td>−40, 2, 32</td>
<td>6.94</td>
<td>a</td>
</tr>
<tr>
<td>Left triangular part of inferior frontal gyrus</td>
<td>−46, 42, 4</td>
<td>6.35</td>
<td>44</td>
</tr>
<tr>
<td>Left insula</td>
<td>−32, 16, 8</td>
<td>8.64</td>
<td>322</td>
</tr>
<tr>
<td>Right insula</td>
<td>40, 20, −4</td>
<td>8.80</td>
<td>284</td>
</tr>
<tr>
<td>Left medial superior frontal gyrus</td>
<td>−4, 26, 38</td>
<td>7.36</td>
<td>157^b</td>
</tr>
<tr>
<td>Right medial superior frontal gyrus</td>
<td>4, 30, 42</td>
<td>5.62</td>
<td>b</td>
</tr>
<tr>
<td>Left superior frontal gyrus (SMA)</td>
<td>0, 18, 54</td>
<td>6.04</td>
<td>63</td>
</tr>
<tr>
<td>Right middle frontal gyrus</td>
<td>−28, 60, 8</td>
<td>6.25</td>
<td>41</td>
</tr>
<tr>
<td>Right middle frontal gyrus</td>
<td>50, 26, 30</td>
<td>5.40</td>
<td>8</td>
</tr>
<tr>
<td>Left paracingulate</td>
<td>−2, −24, 30</td>
<td>5.83</td>
<td>47</td>
</tr>
<tr>
<td>Right anterior medial cingulate cortex</td>
<td>4, 38, 28</td>
<td>6.34</td>
<td>23</td>
</tr>
<tr>
<td>Left caudate nucleus</td>
<td>−14, 8, 12</td>
<td>5.92</td>
<td>6</td>
</tr>
<tr>
<td>Left inferior parietal gyrus</td>
<td>−42, −50, 52</td>
<td>7.06</td>
<td>241</td>
</tr>
<tr>
<td>Left supramarginal gyrus</td>
<td>−48, −40, 28</td>
<td>5.92</td>
<td>4</td>
</tr>
<tr>
<td>Left superior temporal gyrus</td>
<td>−58, −10, 0</td>
<td>5.72</td>
<td>8</td>
</tr>
<tr>
<td>Left hippocampus</td>
<td>−26, −42, 0</td>
<td>5.55</td>
<td>4</td>
</tr>
<tr>
<td>Left cuneus</td>
<td>−8, −84, 26</td>
<td>6.17</td>
<td>21</td>
</tr>
<tr>
<td>Right medial occipital lobe</td>
<td>12, −60, 8</td>
<td>6.01</td>
<td>37</td>
</tr>
<tr>
<td>Right cerebellium</td>
<td>2, −62, −26</td>
<td>5.98</td>
<td>65</td>
</tr>
</tbody>
</table>

**Table 3**

Brain areas showing an effect of learning.

<table>
<thead>
<tr>
<th>Structure</th>
<th>(x, y, z)</th>
<th>T-value</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Situation-based learning effect ((S + T) (-) (S)) masked by (S + T) (-) (S))</td>
<td>−48, 14, 28</td>
<td>6.35</td>
<td>148^c</td>
</tr>
<tr>
<td>Right supramarginal gyrus</td>
<td>50, 20, 14</td>
<td>6.67</td>
<td>30</td>
</tr>
<tr>
<td>Text-based learning effect</td>
<td>−58, −10, 0</td>
<td>5.72</td>
<td>8</td>
</tr>
<tr>
<td>Left middle frontal gyrus</td>
<td>−32, 48, 8</td>
<td>4.44</td>
<td>26</td>
</tr>
</tbody>
</table>

For each area, the coordinates \((x, y, z)\) of the activation peak in the MNI space, peak T-value, and size of the activated cluster in a number \((k)\) of voxels \((2 × 2 × 2\) mm\(^3\)) are shown for all subjects \((n = 31)\).

**Figure 3.** Comparison of the learning effect on brain activity under the situation-based learning (SL words: \(S + T\) and \(S\)) and text-based learning (TL words: \(S + T\) and \(S\)) conditions. Activation in (a) the right supramarginal gyrus \((x, y, z; 60, −20, 18)\) for SL words and activation in (b) the left middle frontal gyrus \((x, y, z; −34, 48, 8)\) for TL words were observed. Activation was set at threshold of \(P < 0.05\) using a family-wise error rate. The activation profile of each area represents the parameter estimates in each condition \((S + T, S, T, −S, T, −S)\), and \((T)\) relative to baseline. Error bars represent the standard error.

**Discussion**

Two major findings emerged from the present study. First, the brain processes L2 words differently depending upon how they were learned. The right supramarginal gyrus was involved in the retrieval of L2 words that were encoded in a social context, whereas the left middle frontal area was involved in the retrieval of words encoded through translation of their L1 counterparts. Second, the left triangular part of the IFG was recruited when words that were encoded under the situation- or text-based condition were retrieved using the opposite condition.

**Effect of situation-based learning**

The right supramarginal gyrus may represent the multimodal information that comprises the action contexts associated with each SL word, which was acquired through the abstraction of the functional meaning of a word (i.e., learning how to use the words in the social context) during the learning session. During situation-based learning, the subjects watched several events consisting of multiple physical and social components that had multimodal real-life representations,
including individuals’ behaviors and details in the environment, and combined significant elements to understand the word’s meaning. Functionally, the right parietal lobe, including the supramarginal gyrus, has long been recognized as an association cortex that integrates multisensory information such as space, motion, and sound (Macaluso and Driver, 2003). Thus, the supramarginal gyrus may be activated when people integrate complex cognitive processes including observing and simulating a motor action (Chong et al., 2008b; Grezes and Decety, 2001), reading the intention of an agent performing an action (Fogassi et al., 2005), imitating actions (Chaminade et al., 2005), and processing motor imagery (Danckert et al., 2002).

Our data suggest that situation-based learning may be mediated by the mirror neuron system. The right supramarginal gyrus is involved in the perception of purposeful actions and is considered one of the constituents of the mirror neuron system (Chong et al., 2008a; Fogassi et al., 2005). Mirror neurons have been proposed to play an important role in imitation learning and language acquisition by observing others (Arbib, 2005; Gallese, 2008). When acquiring a language, learners likely rely on associations between the movements of the body and the context in which the words are spoken (Tomasello and Akhtar, 2005). Thus, the present findings suggest that parietal mirror neurons may be involved in the acquisition of the functional meaning of words through observation of an action and its intention in the social context.

The right supramarginal gyrus might have been activated simply because the subjects observed actors’ actions during encoding (a trajectory effect at the perceptual level) rather than because they learned new words; however, this was not the case. In the present experiment, L1 words learned in a social context also elicited greater activity in the right supramarginal gyrus than did L2 words encoded from text (Fig. 3a). Our findings suggest that a similar cognitive process underlies L2 and L1 word acquisition in a social context.

**Effect of text-based learning**

The retrieval of TL words induced greater activation of the left middle frontal regions than did that of SL words. Activation of the left middle frontal regions may be involved in working memory for rote associative memorization between a target word and its L1 counterpart. The left middle frontal gyrus may be involved in the manipulation of information in the active working memory (Fletcher and Henson, 2001). Additionally, activity in the left prefrontal region increases during word-paired associate encoding (Dolan and Fletcher, 1997), and the middle frontal gyrus is involved in the formation of verbal relational associative learning (Fletcher and Henson, 2001; Woodward et al., 2006). Only verbal information was provided to the subjects in the TL condition; therefore, associating the meaning of a target word with its L1 counterpart likely recruited the left middle frontal gyrus.

Although the activation profile of the left middle frontal gyrus appears to show deactivation for the SL words, it is likely to be an artifact of the baseline activity. This region is a part of a network that shows some degree of activity during the conscious resting state and deactivation during some attention demanding tasks (Raichle et al., 2001; Shulman et al., 1997).

**Inferior frontal gyrus activation in the learning and testing-mismatched condition**

We found that word retrieval under the learning- and testing-mismatched condition evoked greater activation of the left triangular part of IFG than did that under the matched condition. This finding is consistent with our hypothesis that the IFG was associated with the flexible application of stored L2 vocabulary. The IFG has been suggested to support flexible behavior and general cognitive control in various domains (Anderson et al., 2004; Badre and Wagner, 2007; Miller and Cohen, 2001; Thompson-Schill et al., 1997). In particular, IFG activity increases when a person is required to resolve a conflict

---

**Table 4**

<table>
<thead>
<tr>
<th>Interaction effect (mismatched condition vs. matched condition)</th>
<th>Structure</th>
<th>x, y, z</th>
<th>T</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>⎛nS+uT⎞−⎜uS+nT⎟ = ⎜nS−uS⎞+uT−nT</td>
<td>Left triangular part of the inferior frontal gyrus</td>
<td>−52, 20, 14</td>
<td>5.73</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>−50, 14, 28</td>
<td></td>
<td>5.11</td>
<td></td>
</tr>
</tbody>
</table>

For each area, the coordinates (x y z) of the activation peak in the MNI space, peak T-value, and size of the activated cluster in a number (k) of voxels (2×2×2 mm³) are shown for all subjects (n = 31).
between different task sets and rules (Badre and Wagner, 2006; Braver et al., 2003; Huijbbers et al., 2009). The cognitive operations used to encode a word differ from those used to retrieve it under the mismatched condition; thus, the subjects may have required more controlled access to stored L2 information to retrieve words encoded in another context. As Kroll and Stewart (1994) argue, words encoded from different learning contexts may represent different conceptual associations between word phonemes and their meanings. A word phoneme encoded in a social situation has a greater association with its functional concept (i.e., usage), whereas that encoded from text is directly connected to the L1 translation. Thus, subjects under the mismatched condition may have needed to resolve conceptual conflict to appropriately transfer the retrieved conceptual information to the currently relevant task.

The five situations in the movie clips used in the situation-based learning session may comprise five different contexts that were associated with each SL word, in contrast to each TL word that was associated with a single L1 translation across the movie clips. This conceptualization seems to provide a better explanation for the activation profile of the IFG as well as the behavioral data. The abundance of contexts associated with the SL words may afford efficacy in comprehending the word when it is used in a new context or situation. This prediction was supported by the behavioral data (Table 1); the SL words were relatively well understood in both text- and situation-based tests (both are new contexts), while the comprehension of the TL words dramatically deteriorated in the situation-based test (new context), in contrast to the excellent performance in the text-based test (learned context). The high cognitive cost of processing the TL words in a new context is reflected in prominent activation of the left IFG in the $^{15}$O condition, in contrast to the modest degree of activation for the SL words in both the test types (Fig. 4).

Brain areas involved in retrieval of learned L2 words

Several brain regions were more highly activated when subjects tried to retrieve L2 words compared to L1 words. They included the left inferior, middle and superior frontal areas, anterior cingulate cortex, temporal areas, parietal lobe, bilateral insula, and basal ganglia. These results are consistent with previous neuroimaging studies of the acquisition of L2 words that recruited multi-element cognitive processing (Breitenstein et al., 2005; Mestres-Misse et al., 2008; Raboyeau et al., 2004). The left frontal areas are involved in executive aspects of linguistic information during learning and remembering (Gabrieli et al., 1998). The left temporal regions, including the hippocampus, are involved in semantic memory, which is related to storing the meaning of a word (Breitenstein et al., 2005). The left parietal area is associated with phonological storage (Paulesu et al., 1993), and the anterior cingulate cortex is associated with attention during verbal task (Raboyeau et al., 2004). The insula and subcortical areas are involved in the control of articulation, an important function for learning the phonetics of new word (Callan et al., 2003; Chee et al., 2004).

Implications, limitations, and further studies

The present study has several methodological limitations. First, our target words included verbs, adjectives, and greetings that may have been easily learned in a social context and generalizations to the learning of different types and categories of L2 words, such as abstract words, should be made with caution. Second, the experiment was conducted in one day, although the cortical mechanisms involved in long-term memory of L2 words were demonstrated. Had testing been conducted after a week or a month, different cognitive processes relating to consolidation and restructuring of semantic memory would have come into play. Future studies are needed to examine the long-term effects of encoding on cortical processing. Despite these limitations, our study provides neurological evidence that the type of word learning has an impact on brain networks during retrieval.

Acknowledgments

This research was supported in part by JST/RISTEX 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). M.S. is supported by a Grant-in-Aid for Young Scientists (KAKENHI 18680026) from MEXT.

References