Knowledge-based Correction of Flash-lag Illusion

Yasuki Noguchi[1,2] and Ryusuke Kakigi[1,3]

Abstract

■ A flash is perceived to lag spatially behind a moving object even when the two retinal images are physically aligned (flash-lag effect, FLE). Here we show that this robust illusion can be diminished by a knowledge of letters in the observer's brain. When moving and flashed segments in the FLE made the shape of a Kanji letter (ideographic characters used in Japan), the magnitude of the illusory lag perceived by Japanese subjects was significantly reduced compared to when conventional geometric (nonletter) segments were used. This diminishment was not observed when a pseudo-Kanji letter was presented to Japanese subjects or when non-Japanese English-speakers (who do not have a knowledge of Kanji) saw a real Kanji letter, indicating that the reduction in the FLE was induced by a retrieval of the knowledge (shapes of letters) stored in the observer's brain. Furthermore, measurements of neural activities by magnetoencephalography showed that the initial brain response, in which the effect of the knowledge became evident, occurred as early as 160 msec after the appearance of the flashed segment. These results demonstrated a substantial influence of knowledge on the flash-lag illusion and further suggest a rapid response of the knowledge-based perceptual pathway in the human brain. ■

INTRODUCTION

When a brief flash is presented physically aligned with a continuously moving object, the flash appears to lag behind the moving object (Nijhawan, 1994; Mackay, 1958) (Figure 1A), a phenomenon called the flash-lag effect (FLE). Since interest in this areas was renewed by Nijhawan (1994), many theories have been proposed to explain this phenomenon, including motion extrapolation (Nijhawan, 1994), differential latency (Purushothaman, Patel, Bedell, & Ogmen, 1998; Whitney & Murakami, 1998), attention (Baldo & Klein, 1995), temporal integration (Krekelberg & Lappe, 2001), and postdiction (Eagleman & Sejnowski, 2000). The precise neural mechanisms underlying the FLE, however, remain unclear (Nijhawan, 2002).

Recently, Watanabe, Nijhawan, Khurana, and Shimojo (2001) reported a significant influence of positional relationships between the moving and flashed segments on the magnitude of the FLE. When a flash was presented near the leading position of a moving stimulus, subjects perceived a larger FLE compared to when it was presented near the trailing position. If subjective positions of the leading and trailing edges of the moving object were changed by manipulating Gestalt factors of constituents, the FLE's magnitude was found to vary in accordance with the manipulations. Because those effects of positional relationships suggest an involvement of high-level visual processing in the FLE, they argued that a single low-level visual process, such as shorter visual latency for motion signals and longer visible persistence for a flashed item, cannot alone account for the various FLE phenomena.

In the present study, we extended the finding of Watanabe et al. (2001) and examined whether ''knowledge'' plays a significant role in determining the magnitude of the FLE. Although Watanabe et al. investigated the relationship between the FLE and the geometric (positional) configurations of the flashed and moving segments given as bottom-up inputs, here we tested a situation where the configurations (that would interact with the FLE) were shaped in a top-down manner. Specifically, we utilized the linguistic knowledge (shapes of letters) stored in the observers' brain. In the experimental conditions of our study, two segments constructing a Kanji character (ideographic characters used in Japan) were presented to Japanese subjects, one as the moving object and the other as the flashed object in the FLE (Figure 1B, KC stimulus). In such a situation, the FLE would produce a ''lagged'' percept of the moving and flashed segments, whereas a knowledge of the letter would indicate a ''nonlagged'' image because of its familiarity or congruency with the canonical template (knowledge) in the brain. Therefore, if knowledge has a significant influence on the FLE, a successful retrieval of knowledge would lead to a cancellation or reduction of the flash-lag illusion compared to when conventional (nonletter) segments are presented.

In addition to these behavioral observations, we further investigated neural activities during this competition between knowledge and the illusion. Of special interest

[1]National Institute for Physiological Sciences, Myodaiji, Okazaki, Japan, [2]Nagoya University, Nagoya, Japan, [3]RISTEX, Japan Science and Technology Agency, Japan

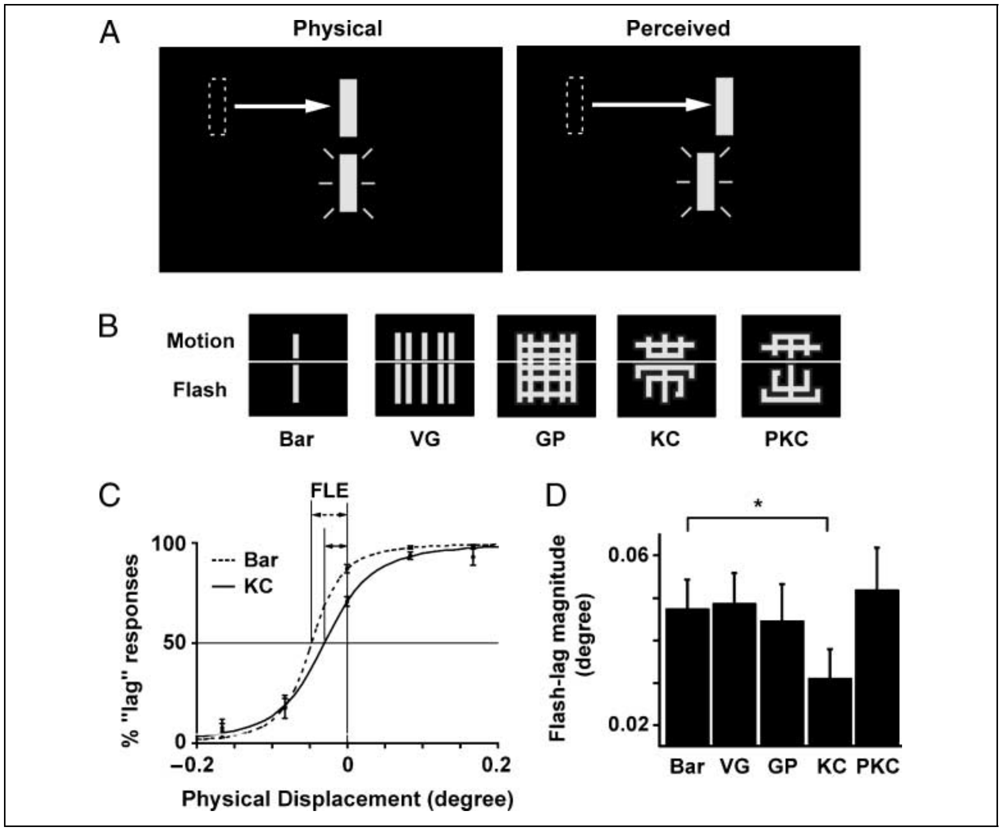

Figure 1. Stimuli and results in Experiment 1. (A) Linear display of the flash-lag effect (FLE). (B) Stimulus set in Experiment 1 using Japanese subjects. In each stimulus, the segments above and below the white horizontal line corresponded to the moving and flashed segments, respectively. VG = vertical grating; GP = grid pattern; KC = Kanji character; PKC = pseudo-Kanji character (made by flipping vertically each segment of the KC). (C) Psychometric functions averaged across eight Japanese subjects in Experiment 1. The horizontal axis indicates the physical displacement between the two segments (positive: the flash lagged behind the motion). For each displacement of the Bar (dotted line) and KC (solid line) conditions, the percentage that subjects answered "lag" (mean ± *SE* across subjects) was plotted and fitted with a nonlinear function (see Methods). The magnitude of the FLE was estimated as the shift of the point of subjective equality (PSE, the points where the psychometric curve crosses the 50% threshold) from zero. (D) Magnitude (mean ± *SE* across subjects) of the FLE perceived by Japanese subjects. $*p < .05$.

was a temporal profile of the knowledge-based pathway in the brain. Previous neuroimaging studies have reported a significant activation of sensory brain regions when perceptual knowledge was recollected (Goldberg, Perfetti, & Schneider, 2006; Martin & Chao, 2001; Mummery, Patterson, Hodges, & Price, 1998). According to recent studies, this activation is elicited in a modality-specific manner [i.e., the visual areas are activated when visual knowledge (e.g., the color or shape of an object) is retrieved, whereas the recollection of gustatory knowledge (e.g., taste) induces activation in the gustatory areas, etc.] (Goldberg et al., 2006), and damage in the fusiform gyrus actually impairs the recollection of visual knowledge (Vandenbulcke, Peeters, Fannes, & Vandenberghe, 2006). However, the temporal dynamics of the knowledge-based pathway has remained unclear, due to the limited temporal resolution of previous approaches using positron emission tomography (PET) or functional magnetic resonance imaging (fMRI). How rapidly is perceptual knowledge retrieved and applied to the visual percept? We therefore examined this issue using magnetoencephalography (MEG) that has a high temporal resolution (order of millisecond) and a moderate spatial resolution. Specifically, we compared the visual-evoked fields (VEFs) between two types of trials in which the effect of perceptual knowledge was strong and not strong, with the bottom-up visual input held constant. The dissociation timing of these two VEFs would provide a temporal index of when the knowledge-based pathway is activated and would begin to affect the formation of visual percepts.

METHODS

Subjects

The main part of the present study consisted of four experiments: two behavioral experiments using Japanese subjects (Experiments 1 and 2), a behavioral experiment using non-Japanese subjects (Experiment 3), and an MEG experiment using Japanese subjects (Experiment 4). Eight (6 men and 2 women) and five (3 men and 2 women) native Japanese speakers participated in Experiments 1 and 2, respectively. In contrast, all nine participants in Experiment 3 were non-Japanese native English-speakers (4 men and 5 women; 2 from the United States, 1 from Canada, 4 from Australia, and 2 from New Zealand; average duration of stay in Japan: 10.1 months). Finally, we conducted the MEG session using seven out of the eight Japanese subjects in Experiment 1. All subjects had normal or corrected-to-normal visual acuity and were naive to the purpose of the experiments. Informed consent was received from each participant after the nature of the study had been explained. Approval for these experiments was obtained from the ethics committee of the National Institute for Physiological Sciences, Okazaki, Japan.

Stimuli and Task

In all four experiments, we used a conventional linear motion display (Krekelberg & Lappe, 2001) to make a flash-lag illusion (Figure 1A). In the standard condition (Bar), a rectangular bar segment ($0.375^\circ \times 0.083^\circ$, 43 cd/m^2) appeared at the left or right edge of a black screen (0.3 cd/m^2 in Experiments 1 to 3 and 0.7 cd/m^2 in Experiment 4) and moved horizontally at a speed of 5 deg/s. The direction of motion was from left to right in one half of the trials (when the bar appeared at the left edge), and from right to left in the other half (when the bar appeared at the right edge). At the mid-trajectory of the bar's motion, a second bar ($0.583^\circ \times 0.083^\circ$, 43 cd/m^2) was flashed below the moving bar for 1 frame (16.7 msec). There was an interval of at least 417 msec from the onset of motion of the upper bar to the flashing of the lower bar. The minimum distance between the red crosshair and the lower end of the second bar was 1.46°. Subjects were required to attend to the positional relationships between the two central bars at the moment of the flash and judge whether the upper (moving) bar was located to the left or right of the lower (flashed) bar (forced-choice). They were also instructed to neglect the shape combinations between two segments. The position of the flash was variable across trials so that the subjects could not perform the task on the basis of the absolute (not relative) position of the moving segment at the moment of the flash.

In Experiment 1 (Japanese subjects), we tested the five types of morphological combinations in Figure 1B: the conventional two-bar segments (Bar), vertical grating (VG), grid pattern (GP), Kanji character (KC, a letter meaning ''belt''), and pseudo-Kanji character (PKC) (note that all stimuli have a pair of vertical bars at their central positions). We used VG and GP to see whether the geometric coherence between two segments has a significant influence on the perception of their spatial relationship. On the other hand, KC was employed to test the possibility of a knowledge-based cancellation of illusory percept. The PKC stimulus was composed of the same local features as the KC, but has no meaning and is not used in everyday life, thus controlling for the complexity and large visual size of the KC letter compared to the Bar stimulus. Trials of these five conditions were randomly intermixed and subjects were required to judge the relative position of the upper (moving) compared to the lower (flashed) bar (left or right), irrespective of morphological combinations of two segments. One experiment contained 112 trials of the Bar and PKC conditions, and 56 trials of the remaining three conditions. To estimate the size of the illusory lag perceived by the subjects, the magnitude of the physical lag was manipulated and randomly set at one of five points ranging from -0.167° to 0.167° (positive: the flash lags behind the motion).

In Experiment 2, we tested the reproducibility of the previous experiment using another Kanji letter (meaning ''thunder''; Figure 2A). Only the KC and PKC conditions (224 trials for each) were included in this experiment and lag magnitudes in the two conditions were directly compared. The magnitude of the physical lag was randomly set at one of nine (not five) points within -0.167° to 0.167°, and the speed of the moving segment was changed from 5 to 2.5 deg/s in order to confirm the speed-proportional characteristic of the FLE (Krekelberg & Lappe, 2001). Other procedures were identical to Experiment 1.

In Experiment 3, we investigated the effect of the Kanji configuration on the FLE using non-Japanese subjects who had no knowledge of Kanji letters. Before the experiment, the proficiency in Kanji characters of all English speakers was carefully checked using the Kanji test shown in Figure 3. Because the Kanji to be used in the experiment was of the level for fourth grade elementary students (9 or 10 years old), we selected 20 Kanji letters from this level for the test. Each subject was required to give the pronunciation and meaning of these characters, and the number of correct answers was defined as the scores (pronunciation: 20, meaning: 20, total: 40) of that subject. We then conducted a flash-lag experiment using the Bar, KC, and PKC conditions on the basis of the same procedures as for Experiment 1. Furthermore, to investigate the knowledge-based correction of the FLE in these subjects, two conditions were newly introduced: letters of the alphabet (AL) and pseudo-alphabet (PAL). As shown in Figure 4A, the AL stimulus was either ''t'' or ''f'' in lower case whereas the PAL stimuli were mirror-reversed images of those letters. For each of these five conditions, 88 trials were given to each subject (the AL and PAL conditions were not tested in Japanese subjects in Experiments 1 and 2 because most Japanese are exposed to letters of the alphabet from a young age).

In the final MEG session using Japanese subjects again, we focused on the three types of stimuli (Bar,

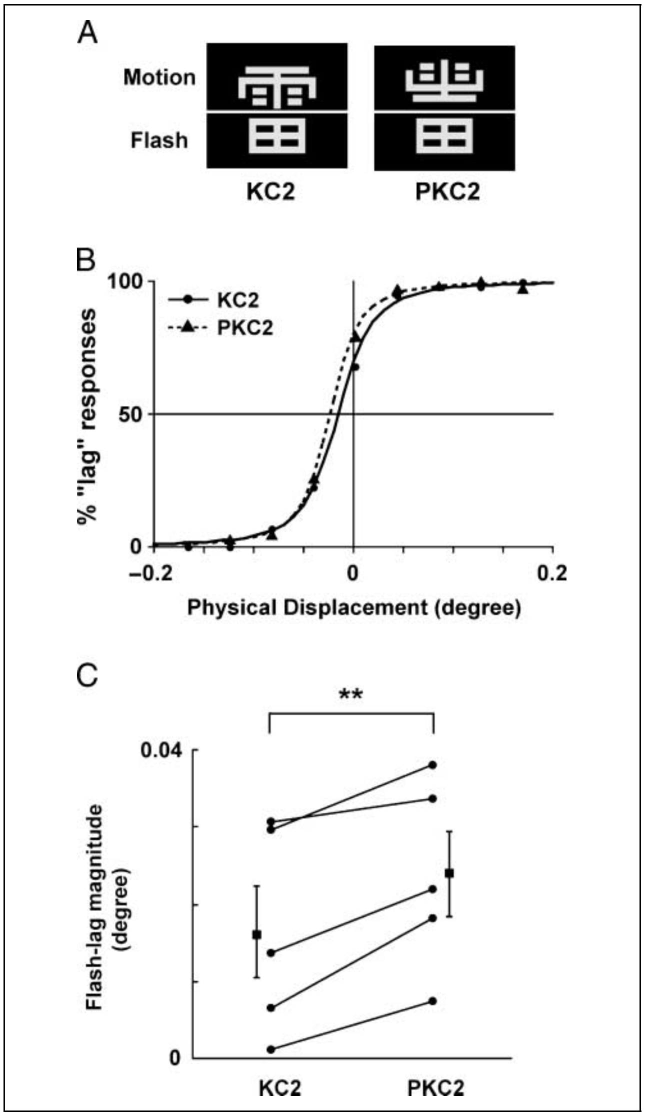

Figure 2. Results in Experiment 2 in which another Kanji character (KC2) was presented to Japanese subjects. (A) Stimulus set. Only the KC2 and PKC2 conditions were tested and directly compared. (B) Psychometric functions averaged across the five subjects. Note that the speed of the moving segment was reduced by half (2.5 deg/s) compared to Experiment 1 (5 deg/s), producing an overall decrease in the lag's magnitude. (C) Magnitude of the illusory lag. Each line shows the data for each subject. Means and *SE*s across the subjects are represented by square points and error bars, respectively. $^{**}p < .01$.

KC, and PKC) to acquire reliable VEFs for each condition. Neural activities in response to the appearance of the flashed segment (that would activate in the KC condition the knowledge-based pathway on being combined with the moving segment) were measured and averaged as VEFs. The 248 trials were tested for each stimulus and behavioral data in the MEG session were pooled with those in Experiment 1. As in the other three experiments above, trials of all conditions were randomly intermixed. Other details were identical to Experiment 1.

Psychophysical Estimation of the FLE's Magnitude

We estimated the magnitude of the illusory lag by calculating the point of subjective equality (PSE) at which subjects were equally likely to say that the flashed segment appeared to ''lead'' and ''lag'' the moving segment. For each condition of each subject, the probability that he or she answered ''lag'' (% lag) was plotted for each physical displacement between the two segments (Figure 1C). The magnitude of the physical lag ranged from −0.167° to 0.167° in all experiments. Conforming to a previous study on the FLE (Kanai, Sheth, & Shimojo, 2004), the % lag data were then fitted with a nonlinear psychometric function,

$$F(x) = 0.5 + (a + bx)/2\left[1 + (a + bx)^2\right]^{1/2}$$

where x is the magnitude of the physical lag, and a and b were free parameters estimated with a least squares criterion. The PSE was defined as the crossing point of this curve with the 50% threshold and obtained as

Write the pronunciation and meaning of the Kanji on the left

You may answer by using either Japanese (Kana letters) or English (alphabets).

	漢字 (Kanji)	発音(pronunciation)	意味(meaning)
1	陸	(riku)	(land)
2	腸	(chou)	(bowel)
3	象	(zou/shou)	(elephant)
4	松	(matsu/shou)	(pine tree)
5	軍	(gun)	(army)
6	愛	(ai)	(love)
7	隊	(tai)	(band/team)
8	帯	(obi/tai)	(belt)
9	孫	(mago/son)	(grandchild)
10	毒	(doku)	(poison)
11	胃	(i)	(stomach)
12	害	(gai)	(harm)
13	芽	(me/ga)	(bud)
14	席	(seki)	(seat)
15	賞	(shou)	(award)
16	粉	(kona/fun)	(powder)
17	鏡	(kagami/kyou)	(mirror)
18	量	(ryou)	(quantity)
19	巣	(su/sou)	(nest)
20	底	(soko/tei)	(bottom)

Figure 3. The test used to measure proficiency in Kanji (parentheses: correct answers in English). Subjects were required to give the pronunciation and meaning of 20 Kanji letters selected from the level for fourth-grade elementary students in Japan (9–10 years old). A knowledge of Japanese characters other than Kanji (Hiragana or Katakana) was unnecessary to fill the blanks because subjects could answer all questions using letters and English words.

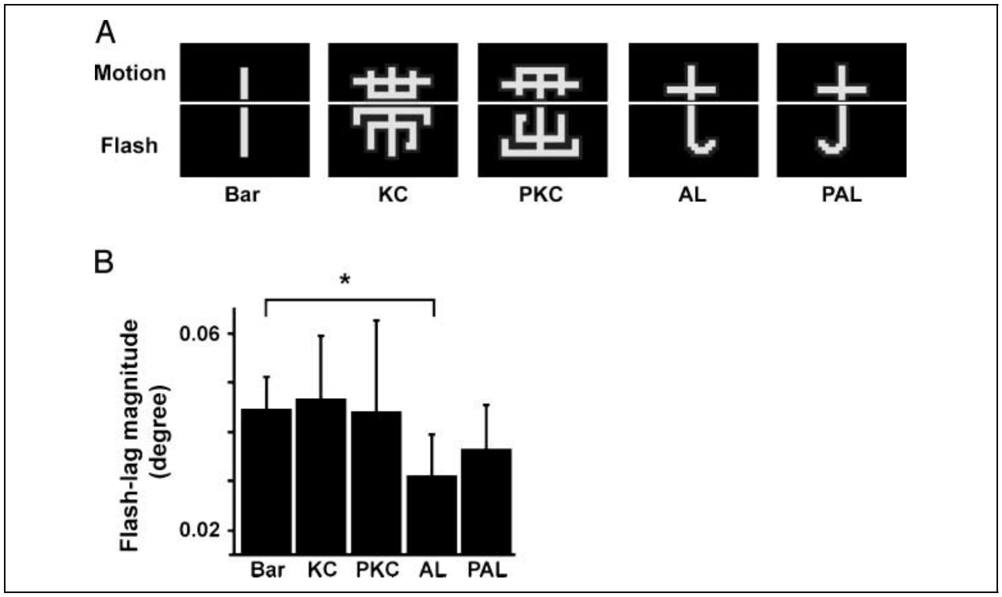

Figure 4. Results in Experiment 3 using non-Japanese native English-speakers. (A) Stimulus set. Instead of VG and GP, letters of the alphabet (AL) and pseudo-alphabet (PAL, made by flipping the AL letters horizontally) were newly employed to test the knowledge-based correction of the FLE in those subjects. (B) The magnitude of the FLE (mean ± *SE* across subjects) is shown in the same format as Figure 1D. Although no Kanji-induced reduction of the FLE was observed in those subjects, the AL condition produced a significant change in the perceived lag magnitude. $*p < .05$.

($-a/b$). Fitting coefficients (r^2) of the psychometric function averaged across all conditions were .98 (range: .87–1) and .99 (range: .98–1) in Experiments 1 and 2, respectively. In Experiment 3, however, fitting coefficients were found to be low in some subjects [mean $r^2 = .92$, range: .29–.99, significantly smaller from those in the Japanese group, $t(82) = 2.9$, $p < .01$]. We thus used a linear interpolation instead of the nonlinear fitting in this experiment.

MEG Recordings and Data Analyses

In Experiment 4, VEFs in response to an appearance of the flashed segment in each condition were recorded with a helmet-shaped 306-channel MEG system (Vectorview, ELEKTA Neuromag, Helsinki, Finland), which comprised 102 identical triple sensor elements. Each sensor element consisted of two orthogonal planar gradiometers and one magnetometer coupled to a multi-SQUID. The MEG signals were recorded with 0.1–200 Hz band-pass filters and digitized at 600 Hz. Eye positions of subjects were simultaneously monitored using an infrared eye tracker (Iscan Pupil/Corneal Reflection Tracking System, Cambridge, MA). The VEFs were averaged from 100 msec before to 500 msec after the onset of the flashed segment, with the prestimulus period (initial 100 msec) used as a baseline. Epochs in which signal variations were larger than 3000 fT/cm were excluded from the averaging. We conducted two types of selective averaging. First, we compared the VEFs in the KC and PKC conditions to investigate whether neurons in the ventral visual pathway showed letter-selective activation (the stimulus-based averaging; Figure 5B). There was no physical lag in either condition. Subsequently, we contrasted two types of trials within the KC condition: the trials in which the subjects reported ''lag'' and ''no-lag'' (''lead'') percepts (the response-based averaging, Figure 5C). If the Kanji template in the brain was strongly activated, a substantial decrease in the magnitude of the FLE would occur, resulting in an increased number of no-lag trials. In contrast, if the activation of the Kanji template was relatively weak, the proportion of lag trials would be far dominant over that of the no-lag trials. Thus, a comparison of the lag and no-lag trials within the KC condition broadly reflects a contrast between when the activation of Kanji knowledge was weak and strong, respectively. As a control, we also compared the lag and no-lag trials within the PKC condition where the knowledge-based correction was not induced, which produced a total of six VEFs for each subject (KC, PKC, lag and no-lag in KC, and lag and no-lag in PKC).

For each of those VEFs, we conducted a single equivalent current dipole (ECD) estimation to identify the location of the source in the occipito-temporal region (Figure 6). We adopted a spherical head model on the basis of individual MR images (Hamalainen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). After all VEFs were digitally filtered (0.1–30 Hz), the locations of ECDs that explained the distribution of the magnetic fields over at least 20 channels around the signal maxima were estimated using the least squares method. Conforming to the criteria in previous studies (Noguchi, Inui, & Kakigi, 2004; Nishitani & Hari, 2002), we accepted only dipoles that accounted for at least 80% of the field variance at the peak. These ECDs were successfully identified from the right hemisphere in four subjects and from the left hemisphere in the remaining three subjects. Because we compared the KC versus PKC conditions and the lag versus no-lag trials within the KC or PKC condition, six ECDs corresponding to the six VEFs were estimated for each subject. For each ECD

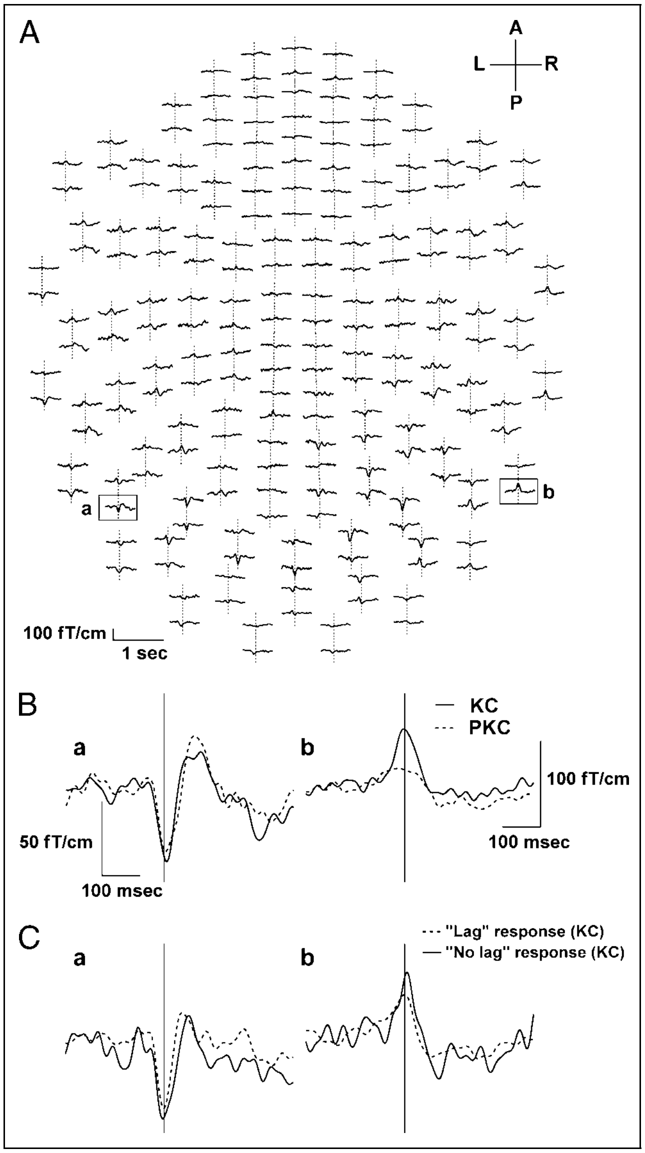

Figure 5. MEG responses of the Japanese subjects during the cancellation of the FLE. VEFs recorded from one representative subject are shown. (A) VEFs in the 204 planar sensors in the KC condition. The dotted vertical lines show 160 msec after the onset of the flashed segment. (B) Comparisons of the VEFs between the KC (solid) and PKC (broken) conditions at the two sensors: a and b (marked in A). Note the clear difference in the magnitude of the 160-msec component, although the KC and PKC stimuli were identical in their local visual features. (C) Same as B, but the lag (broken) and no-lag (solid) trials within the KC condition were compared. Although the physical inputs were identical, deflections of the 160-msec component tended to be larger in the no-lag trials in which the size of the FLE was strongly diminished and the subjects reported "no-lag" percept. Note that, for all data in this and subsequent figures (Figures 6 and 7), there was no physical lag between the moving and flashed segments.

estimated, we then calculated the change of source strength over time, with the position and orientation of the dipole kept fixed. These time series of the ECDs were then averaged across all subjects (Figure 6B–D). The data in the right and left hemispheres were pooled in these figures.

Additional Experiments Using Japanese Subjects

In addition to the main experiments (Experiment 1–4), we conducted three behavioral experiments using Japanese subjects ($n = 5$) to clarify the effect of a knowledge of Kanji on the FLE. The first experiment was designed to examine the possibility of a priming effect between the upper and lower segments of the KC trials. Given the intimate relationships of these two segments (e.g., in the daily lives of Japanese subjects), it can be expected that the presentation of the moving segment of the KC stimulus facilitates the processing of the flashed segment subsequently presented. This may lead to a selective shortening of the neural latency to the flashed segment in the KC (compared to the PKC) condition, resulting in a decrease in the magnitude of the FLE as the latency difference hypothesis predicts (Whitney & Murakami, 1998). To control for this effect, we used a positional judgment task in Figure 8A. In this experiment, the upper segment of the KC or PKC was presented (200–600 msec, variable), followed by the flashing of the lower segment (16.7 msec). The upper segment was static throughout the trial, and the lower segment was identical between the KC and PKC conditions. The lower segment was slightly (~0.167°) displaced relative to the upper segment in either a leftward or rightward direction, and the subjects judged the positional relationship of the two segments as quickly as possible. They pressed one button when the lower segment was located to the right of the upper segment, and another button when it was to the left (forced choice). If there is a priming effect between the two segments of the KC (compared to the PKC) stimulus, the processing of the lower segment should be facilitated by the upper segment of the KC presented in advance, leading to shorter reaction times (RTs) in the KC than PKC condition. An experimental session contained 24 trials (12 for KC and 12 for PCK, randomly intermixed).

The second uncontrolled factor is a temporal binding between the upper and lower segments. Because of the compatibility with knowledge stored in the brain, there may be a stronger temporal binding between the two segments in the KC relative to the PKC stimulus (i.e., the subjects were more likely to perceive the two halves of the KC as co-occurring, producing a worse temporal discrimination). This may also induce the smaller FLE magnitude in the KC than in the PKC condition (Whitney & Murakami, 1998). We examined this possibility using a conventional temporal-order judgment (TOJ) task. If the temporal binding of the two segments is stronger in the KC than in the PKC condition, there should be some difference in the difficulty in performing the TOJ task between the two conditions. As shown in Figure 8B, either the upper or lower half of the KC or PKC letter was presented, followed by the other half. Durations of both the first and second stimuli were 12 msec (we

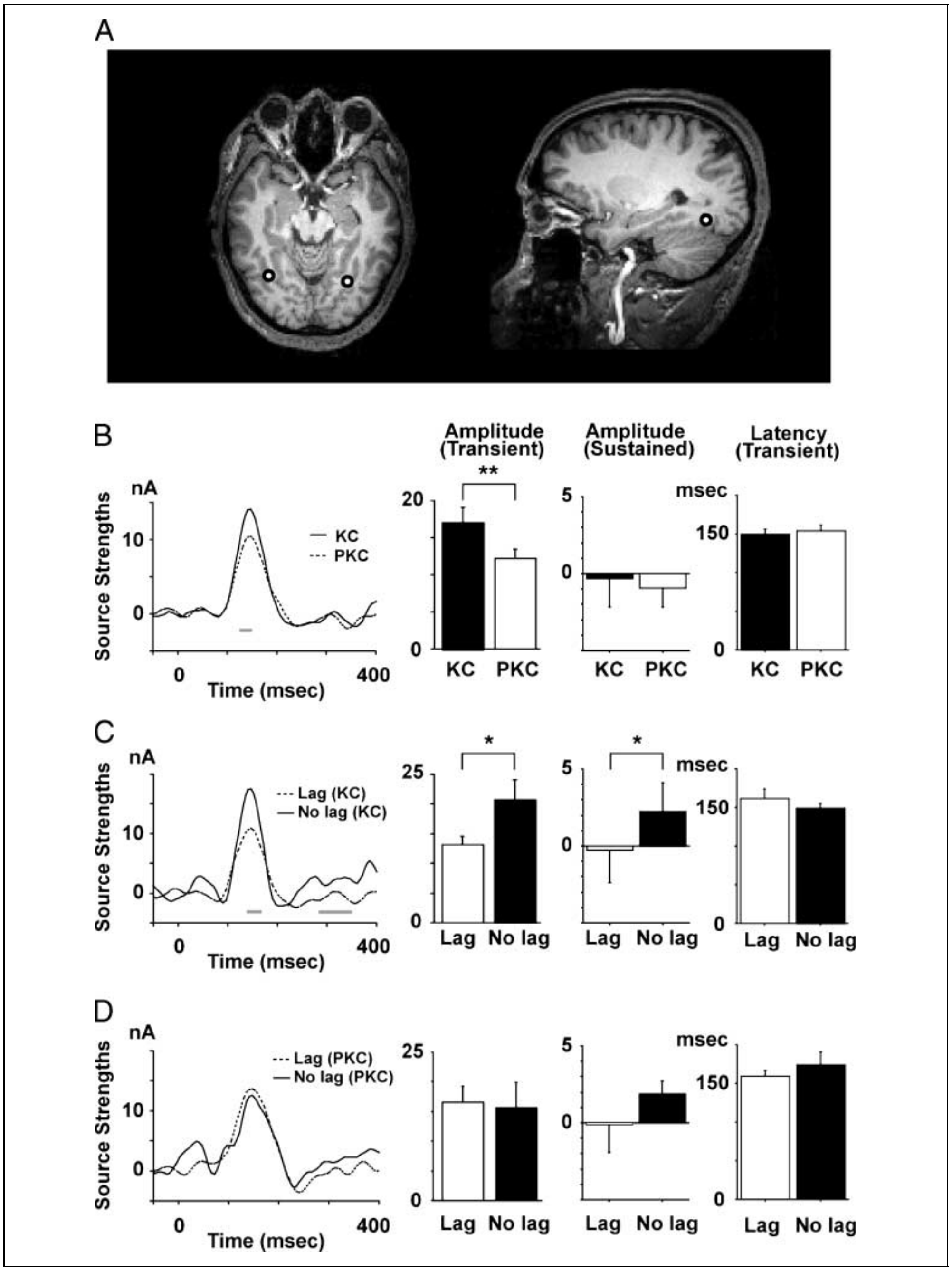

Figure 6. Single ECD analysis applied to the 160-msec component over the left or right occipito-temporal region. (A) Anatomical source locations of the VEFs induced by the flashing of the lower segment under the KC and PKC conditions. (B) Comparisons of neural activities (the ECD source strength over time) between the KC and PKC conditions. The four panels indicate (from the left) the neural time series, peak amplitudes of the 160-msec component (transient activity), mean amplitudes in 300–350 msec (sustained activity), and peak latencies of the 160-msec component (time zero indicates the onset of the flashed segment). The gray line below the time series (leftmost panel) indicates the time range in which two time series were significantly dissociated ($p < .05$, point-by-point paired t tests), and the error bars in the bar graphs denote *SE*s across the subjects. The ventral visual region showed stronger neural responses to the KC than PKC stimulus at a latency of 160 msec. $^{**}p < .01$. (C) Comparison of the "lag" and "no-lag" trials within the KC condition. Both transient and sustained activities became larger when the illusory lag was strongly reduced by the KC configuration ("no-lag" response). $^{*}p < .05$. (D) Same as C, but a comparison of the "lag" and "no-lag" trials within the PKC condition where no Kanji-based diminishment of the FLE was induced.

used an 85-Hz CRT monitor in this experiment). The stimulus-onset asynchrony (SOA) between the first and second halves was −24, −12, 0, 12, or 24 msec (positive: lower first), and subjects answered whether the upper or lower half was presented first. One experiment contained 120 trials in which the KC and PKC trials were randomly intermixed.

Finally, in order to examine whether the spatial grouping of the two segments (based on Kanji knowledge) is responsible for the modulation of the FLE's magnitude, we tested the stimulus configurations shown in Figure 8D. In this experiment, the moving segment alone was a real or pseudo-Kanji letter, and the flash was always the simple bar. As in the main experiments, subjects judged the positional relationship between the two segments in a forced-choice manner. Two types of trials (64 for KC and 64 for PKC) were randomly intermixed. If the knowledge-based grouping of the upper and lower segments is critical to the FLE's modulation, no diminishment in the size of the FLE should be expected in this experiment because the flashed segment now has no reason to be grouped with the moving segment in either condition. On the other hand, if the use of the real Kanji shape is critical for the diminishment of the FLE, a selective decrease in the size of the FLE in the KC (compared to PKC) condition should be observed.

RESULTS

Knowledge-based Cancellation of Illusory Lag

Figure 1D shows the magnitude of the illusory lags in the five conditions of Experiment 1. We initially examined an effect of the geometric coherence between the two segments on the FLE, using the Bar, VG, and GP conditions. A repeated-measures one-way analysis of variance (ANOVA) indicated no main effect [$F(2, 14) = 0.10$, $p > .9$] among the three conditions, suggesting that the geometric coherence had no power to modulate the size of the FLE. In contrast, we found a substantial influence of knowledge on the FLE, as revealed by a significant main effect of an ANOVA including the Bar, KC, and PKC conditions [$F(2, 14) = 4.08$, $p < .05$]. Post hoc tests with the Bonferroni correction revealed a significant difference between Bar and KC ($p < .05$), whereas the difference between KC and PKC did not reach significance ($p = .06$). These results overall indicated that the size of the FLE was selectively reduced in the KC condition while being comparable among the other four conditions.

Similar results were obtained in Experiment 2, in which a different Kanji letter was used. Because the speed of the moving segment was reduced from 5 to 2.5 deg/s, the magnitude of the illusory lag was diminished overall compared to Experiment 1 (KC: from 0.031° to 0.016°; PKC: from 0.052° to 0.024°; Figure 2B and C). Although the statistical significance in this comparison (KC vs. PKC) was marginal ($p = .06$) in Experiment 1, we found a lag significantly smaller in the KC than in the PKC condition in this experiment [$t(4) = 5.2$, $p < .01$; Figure 2C]. The results in those two experiments showed that the illusory percept caused by the FLE was partially cancelled out by knowledge of the letters in the observer's brain.

We then conducted Experiment 3 using non-Japanese English-speakers who did not have a knowledge of Kanji letters. The scores in the Kanji test (Figure 3) revealed a negligible knowledge of Kanji in these subjects, showing a clear contrast with the Japanese subjects in Experiments 1 and 2 (mean ± *SE*, non-Japanese: 1.7 ± 0.7 in pronunciation, 1.6 ± 0.7 in meaning, and 3.2 ± 1.3 in total. Japanese: 19.6 ± 0.3 in pronunciation, 19.9 ± 0.1 in meaning, and 39.4 ± 0.4 in total). When those non-Japanese subjects observed the same display as used in Experiment 1, no effect of Kanji knowledge was observed (Figure 4B), as shown by an ANOVA across the Bar, KC, and PKC conditions [$F(2, 16) = 0.02$, $p > .9$]. Instead, a significant main effect was observed among the Bar, AL, and PAL conditions [$F(2, 16) = 6.01$, $p < .05$], and post hoc tests with the Bonferroni correction revealed a significant reduction in the lag in the AL compared to Bar condition ($p < .05$), which paralleled the results in the previous experiments showing the knowledge-based reduction in the lag in Japanese subjects. As described in Methods, the magnitude of the FLE in Experiment 3 was calculated using linear interpolation, rather than nonlinear fitting as in the previous experiments. We confirmed, however, that the main results of Experiment 3 (a significant decrease of illusory lag in the AL condition, but a nonsignificant change in the KC and PKC conditions) were invariant regardless of whether the linear or nonlinear method was used.

Brain Activity at the Moment of the Retrieval of Perceptual Knowledge

In the final experiment, we recorded the VEFs around when the flashed segment appeared and investigated neural activities during the diminishment of the FLE. Infrared recordings of eye position ensured no systematic eye movements during the trials in any of the conditions. Typical MEG waveforms are shown in Figure 5A. Large deflections of neuromagnetic fields were observed 150–160 msec after the flashed segment appeared (vertical dotted lines), centering on the lateral posterior sides of both hemispheres. When we compared the waveforms for the KC and PKC conditions, this 160-msec component was found to be larger for the former (Figure 5B). Moreover, this difference in transient visual activity was also observed in the comparison between the lag and no-lag trials within the KC condition (Figure 5C) in which visual inputs were identical.

Next, we conducted a single ECD analysis of these lateral posterior VEFs. We confirmed that all dipoles were located in the ventral regions around the occipital–temporal junction (see Figure 6A for averaged locations). The mean ± *SE* of coordinates across subjects were (28 ± 2.0, −41 ± 3.6, 60 ± 3.1) for the right hemisphere and (−33 ± 3.7, −38 ± 5.0, 59 ± 1.9) for the left, according to our head-based coordinate system (Noguchi et al., 2004). In this system, the x-axis was fixed with the preauricular points, the positive direction being to the right. The positive y-axis passed though the nasion and the z axis thus pointed upward. There was no significant difference of ECD coordinates between the KC and PKC conditions or between the lag and no-lag conditions ($p > .1$, for all axes of both hemispheres).

Dipole waveforms averaged across all subjects are shown in Figure 6B–D. As the individual data in Figure 5, the comparison between the KC and PKC conditions indicated a significantly larger 160-msec response in the former [$t(6) = 4.7$, $p < .01$; Figure 6B]. These results were consistent with a previous MEG study reporting a letter-selective activation in the occipito-temporal regions at a latency of 160 msec (Parviainen, Helenius, Poskiparta, Niemi, & Salmelin, 2006), and further suggested that the Japanese subjects recognized the KC stimulus as a letter (not just a shape) even in the FLE display (where the shape discrimination was not relevant to the task).

We then compared dipole waveforms in the lag and no-lag trials within the KC condition (Figure 6C), which showed the amplitude of the 160-msec component to be larger in the no-lag than in the lag trials [$t(6) = 2.7$, $p < .05$; middle panel in Figure 6C]. Moreover, the dipole analysis revealed another dissociation between the two VEFs in the later sustained activities around 300 msec [$t(6) = 3.4$, $p < .05$; right panel in Figure 6C]. In contrast, no significant difference was observed in the same comparison within the PKC condition (Figure 6D), where the recollection of visual knowledge was not elicited. Finally, we compared the latency of the 160-msec component between the KC and PKC conditions and the lag and no-lag trials (rightmost panels in Figure 6B–D). No significant difference was observed in any comparisons ($p > .05$ for all).

Possibility of Attentional Modulation of the 160-msec Component

Previous studies on electroencephalography (EEG) have shown that many EEG (and MEG) components could be substantially modulated by attention (Hillyard & Anllo-Vento, 1998; Mangun, 1995). Thus, one possible interpretation of the MEG results above was that the level of attention differed between the lag and no-lag trials, resulting in the larger 160-msec component in the no-lag than in the lag condition. To examine this possibility, we focused on the P1m waveform (80–130 msec) of the MEG data that appeared earlier than the 160-msec component (corresponds to the N1m). Previous studies indicated that the allocation of attention typically increases the amplitude of both the P1m and N1m components simultaneously (Luck, Woodman, & Vogel, 2000; Mangun, 1995). If the difference in the N1m between the lag and no-lag trials (Figure 6) reflected the attentional modulation, the amplitude of the P1m in the present study should also be larger in no-lag than lag trials.

Figure 7 shows comparisons of the P1m and N1m components between the lag versus no-lag trials. In this analysis, we initially selected an MEG sensor where the amplitude of the P1m was maximum (over the 204 planar sensors) in the time frame of 80–130 msec. Consistent with the previous study (Mangun, 1995), these sensors were concentrated near or slightly lateral to the occipital pole (Figure 7A). Figure 7B shows the results in which the waveforms of these sensors were averaged across the subjects. Although the N1m component was significantly larger in the no-lag than in the lag condition [$t(6) = 3.33$, $p < .05$], no difference was observed in the P1m component [$t(6) = 1.88$, $p = .11$]. These results indicated that the modulation of the 160-msec component in Figure 6 did not reflect the attentional level of the subjects.

Spatial Grouping, Not Priming or Temporal Binding

The RTs (mean ± *SE* across subjects) in the first control experiment (the priming effect; Figure 8A) were 376 ±

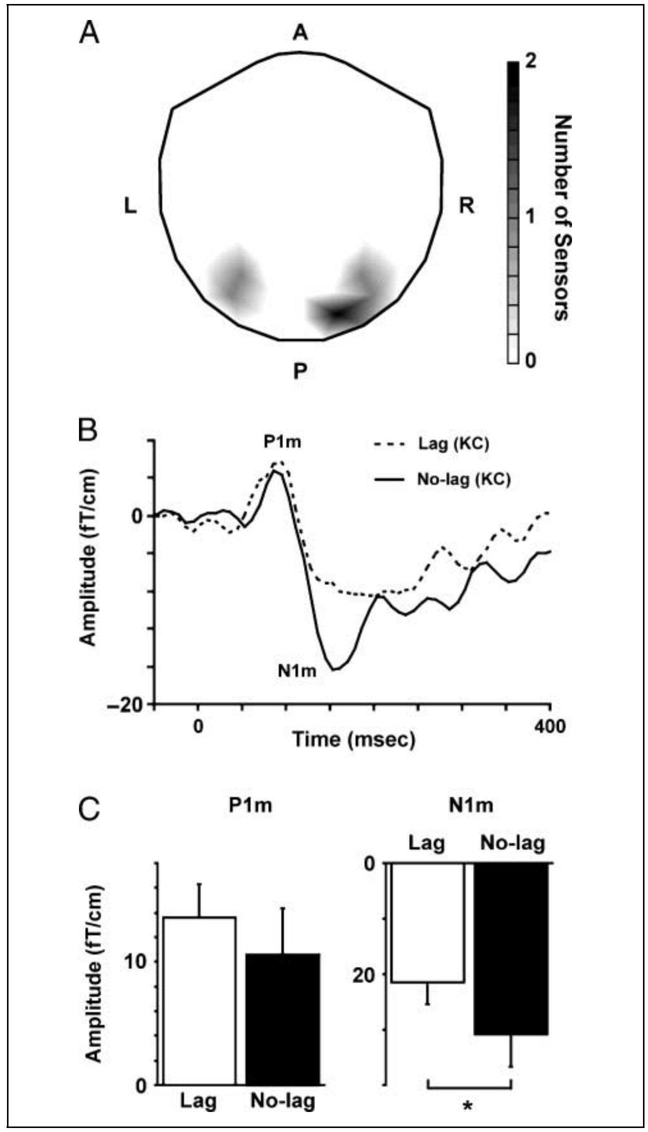

Figure 7. Comparison of the P1m and N1m components. (A) A contour map showing the topographical distribution of the MEG sensors in which the P1m component (80–130 msec) was evidently observed. A = anterior; P = posterior; L = left; R = right. Because we selected one sensor from the data of each subject, a total of seven sensors were used in this analysis. (B) An averaged waveform of the seven sensors selected. The VEFs in the lag (dotted) and no-lag (solid) trials were plotted separately. (C) Mean (±*SE* across subjects) of the peak amplitude of the P1m (left) and N1m (right). A significant difference between the lag and no-lag trials was seen in the N1m but not in the P1m, indicating that the changes in the N1m (160 msec) component in Figure 6 did not reflect the attentional modulation.

19 msec for the KC and 397 ± 26 msec for the PKC trials. No significant difference was observed between those two conditions [$t(4) = 1.38$, $p = .24$], indicating that the priming effect (between the two segments of the KC stimulus) was not the reason for the diminishment of the FLE in Experiments 1 and 2.

In the second control experiment (TOJ, Figure 8B), we plotted the probability that subjects reported the lower segment as the first stimulus, as a function of SOA

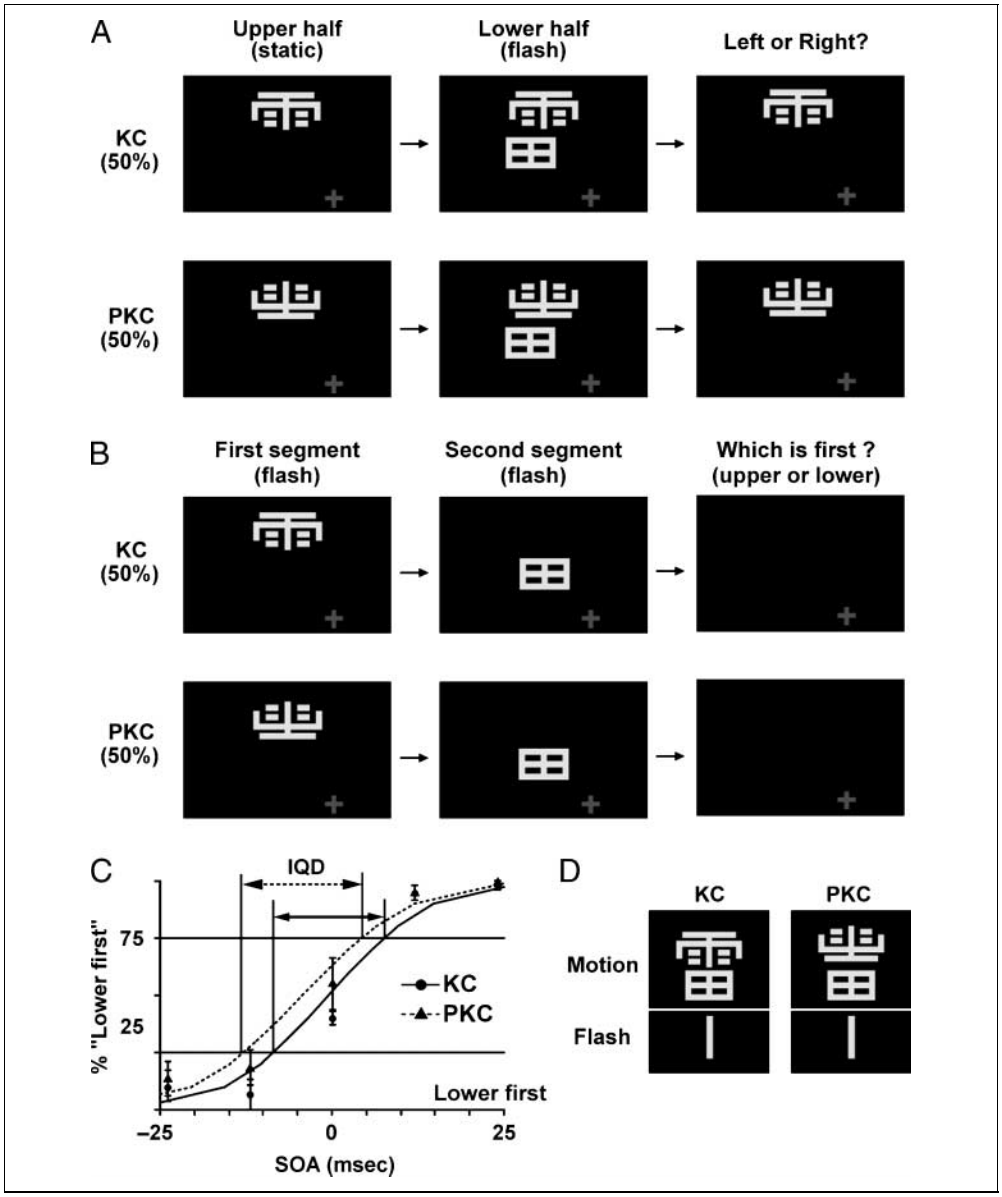

Figure 8. Additional behavioral experiments using Japanese subjects. (A) An experiment to examine a possibility of the priming effect between the upper and lower segments. Following the appearance of the upper segment, the lower segment flashed briefly either to the right or left of the upper segment (displacement: $-0.167°$ to $0.167°$). Subjects reported the relative position of the lower compared to the upper segments as quickly as possible. (B) TOJ task. Subjects judged whether the upper or lower segment was presented first, neglecting the shape (KC or PKC) of the stimuli. (C) The results of the TOJ task. Percentages of ''Lower first'' responses were plotted for each SOA of the two segments (positive: lower first). The SOAs in which subjects answered ''lower first'' in 25% and 75% of the trials were identified by Probit analysis, and the distance between these two points (interquartile distance, IQD) was used as an index of the difficulty of the task. (D) A configuration in which the entire Kanji or pseudo-Kanji was paired with the simple bar as the moving and flashed segments, respectively.

between the two segments (positive: lower first). After fitting the data with a cumulative normal distribution (Probit analysis), we measured the interquartile distance (IQD, a range in the horizontal axis between the 75% and 25% ''lower first'' responses; Figure 8C) for each condition of each subject as an index for the difficulty of temporal discrimination (Leon & Shadlen, 2003). If the temporal binding of the two segments was stronger for the KC than PKC stimulus, this should make the TOJ task more difficult in the KC condition, resulting in a selective increase in the IQD in that condition. The IQD (mean $\pm$ *SE* across subjects) was, however, 16.6 $\pm$ 4.3 msec for KC and 15.9 $\pm$ 3.8 msec for PKC. No significant difference was observed between the two conditions [$t(4) = 0.62$, $p = .57$].

Finally, we investigated the FLE in which the entire Kanji configuration (the moving segment) was paired with the flash of the simple bar (Figure 8D). The mean ($\pm$*SE* across subjects) magnitude of the illusory lag (estimated by the nonlinear fitting above) was $0.033° \pm 0.018°$ for KC and $0.035° \pm 0.016°$ for PKC. There was no significant difference between the two conditions [$t(4) = 0.53$, $p = .62$], indicating that the grouping of the two Kanji halves (based on knowledge) was critical for the diminishment of the FLE.

DISCUSSION

In the present study, we tested whether various morphological combinations of moving and flashed segments in the FLE significantly influence the perceived magnitude of the illusory lag. Consequently, the illusory lag was significantly reduced only when the combinations of two shapes were congruent with the templates of letters stored in the observer's brain, demonstrating an effect of knowledge on the FLE. Moreover, our MEG measurements revealed that the first neural activity in which effect of the knowledge became evident was as early as 160 msec after the flash, suggesting a rapid activation of the knowledge-based pathway in the human visual cortex.

A Factor Critical for the Reduction in the FLE's Size

Although the present results indicate that a knowledge of letters interfered with the illusory lag of the FLE, there might be other explanations for the reduction in the FLE in the KC compared to the Bar condition. First, in addition to the central bars in the Bar stimulus, the KC letter has several pairs of vertical and horizontal bars in the periphery (Figure 1B). Although multiple straight lines are an intrinsic property of Kanji characters, these additional bar pairs might strengthen a linkage between the two segments (Gestalt's laws of ''good continuation'') and attenuate the magnitude of the FLE. One should note, however, that no reduction in magnitude was observed in the VG and GP conditions where stimuli with multiple bar elements were also used. Given the high geometrical coherence in those stimuli, it would be difficult to explain the reduction in illusory lag as caused by the graphical aspects of the KC letters.

A second explanation for the reduction in the FLE is the difference in size and complexity between the Bar and KC stimuli. Because the KC letters are apparently larger and more complex than the Bar stimulus, the spatial attention of observers might be diffused over the entire configuration of the KC letter, which might make the positional judgment of central bars unreliable and attenuate the FLE. Indeed, some previous studies have reported a modulation of the lag's magnitude by spatial attention (Namba & Baldo, 2004; Baldo & Klein, 1995). This account is, however, also unlikely because we observed no reduction in the FLE for the PKC stimulus, which in local visual features and complexity was identical to the KC stimulus. The data for non-Japanese subjects (Figure 4) also provided evidence against this possibility. These results therefore indicate that the congruency between the presented stimuli and the knowledge in the brain is a critical factor for the diminishment of the FLE. As shown in the results of the three additional experiments (Figure 8), the knowledge would facilitate a spatial grouping of the upper and lower segments in the KC stimuli, partially canceling out the illusory displacement produced by the FLE.

Implications and Mechanisms of the Knowledge-based Correction of the FLE

As an extension of a previous study (Watanabe et al., 2001), we showed a significant influence of the global configuration of the two segments on the magnitude of the FLE. A key difference was that the configurations in the present study were based on knowledge, information stored in the brain of the subjects. Our results thus indicate that information other than bottom-up (visual) inputs can be combined with the visual inputs of the FLE to modify the final perception of the stimulus. A recent study using face stimuli (Khurana, Carter, Watanabe, & Nijhawan, 2006) obtained results consistent with ours, although their findings could be explained by possibilities other than knowledge, such as a priming effect and a similarity of visual features between the two halves of their nonchimera faces. Taken together with several studies indicating an involvement of high-level neural mechanisms in the FLE (Kanai & Verstraten, 2006; Nieman, Nijhawan, Khurana, & Shimojo, 2006), our study strengthens the view that at least some portions of the flash-lag illusion cannot be attributed to low-level processing in the visual system.

Several aspects of the present results may be explained by the Bayesian characteristics of the sensory system. Recent studies have indicated that the final outputs in the visual system (percepts) are produced by the probabilistic integration of prior knowledge with the image features provided as inputs (Ernst & Bulthoff, 2004; Kersten, Mamassian, & Yuille, 2004; Ernst & Banks, 2002). Importantly, those priors in perception were found to have plasticity, and thus, can be learned in response to environmental statistics (Adams, Graf, & Ernst, 2004). Given that linguistic knowledge such as letters is usually acquired through repetitive exposure to stimuli, it was highly possible that the Kanji letters implanted in the brain of Japanese subjects worked as the priors of the Bayesian integrations, modifying the final percepts of the FLE in the direction of the internal templates. Indeed, a previous study reported that, in the integration of visual and nonvisual information, the role of nonvisual information becomes important particularly when the visual inputs from an environment are ambiguous (Alais & Burr, 2004). Thus, one sensible interpretation of the present FLE study is that the rapid transient nature of the flash made it hard for the subjects to specify the image at the moment of the flash, resulting in a stronger influence of knowledge on the formation of the percept.

Temporal Dynamics of the Knowledge-based Activation

In addition to those behavioral findings, we investigated the brain activity during the diminishment of the FLE. The responses to the onset of the flash were recorded and compared between the lag and no-lag trials. One possible problem with our MEG measurements would be the moving halves of the KC and PKC stimuli. Because these segments had been presented on the screen before the appearance of the flashed segments, it is possible that the neuromagnetic responses to those moving stimuli caused some distortion of the MEG waveforms in Figures 5 to 7. However, one should note that there was a long interval (at least 417 msec) between the onset of the motion of the upper segment and the flashing of the lower segment. On the other hand, our previous MEG study (Kawakami et al., 2002) showed that large neuromagnetic responses to visual motion at 5 deg/s (equal to the velocity of the moving segment in the present study) were typically seen in the initial 100–200 msec period after motion onset. Thus, it is not likely that the present MEG data for the flashed segment were severely contaminated by the neural signals induced by the moving segment.

As shown in Figure 6C, we found a modulation of the occipito-temporal activity in the KC condition that was closely linked to the behavioral judgments (lag or no-lag) of the subjects. Because the visual inputs were identical between the lag and no-lag trials, this modulation could not be attributed to a difference in bottom-up signals. Given that the same difference was not observed in the lag versus no-lag comparison within the PKC condition (Figure 6D), it is highly possible that the increased neural activity in the no-lag trials of the KC condition reflected the strong activation of Kanji knowledge (Goldberg et al., 2006; Martin & Chao, 2001), and thus, played an important, if not sufficient, role in the activation of the knowledge-based pathway in the brain. This activity tended to be small during the ''lag'' trials (where the knowledge-based correction was relatively unsuccessful) and large during ''no-lag'' trials (where the knowledge-based correction was successful). A novel finding was that this dissociation of the VEFs became evident as early as 160 msec after the appearance of the flashed segment, a latency that has been reported to be in the sensory processing rather than in the later memory or judgment stage (Liu, Harris, & Kanwisher, 2002; Allison, Puce, Spencer, & McCarthy, 1999). Thus, our results support the view that the bottom-up and top-down pathways begin to interact early in the sensory pathway (rather than via slow feedback connections from the higher regions) for the shaping of the final visual percepts. This may be one reason why knowledge-based visual representations [e.g., images of convex or concave patches made through the light-from-above prior (Adams et al., 2004; Sun & Perona, 1998) or hollow-face illusions (Gregory, 1997)] are so strong that they can counter competing bottom-up depth cues and cannot be changed with voluntary attention. It should be noted, however, that our results indicated a rapid pace at the ''beginning'' of the knowledge-based pathway, and thus, do not exclude the contribution of later brain activities to the formation of the percepts or behavioral judgments. Actually, our data indicated a significant dissociation between the lag versus no-lag VEFs in the later period around 300 msec. Although it remains unclear whether the visual knowledge was applied during the period of transient or sustained activity, our data suggest that the strong activity in the ventral regions facilitated the decoding of the knowledge stored in those areas, resulting in competition with the lagged percept produced by the illusion.

Acknowledgments

We thank Dr. S. Shimojo for his helpful comments, Dr. S. Oiwa for her cooperation in the recruitment of non-Japanese subjects, and Mr. O. Nagata and Mr. Y. Takeshima for their technical support. This work was supported by grants from the Japan Society for the Promotion of Science for Young Scientists to Y. N. This study was carried out as a part of ''Ground-based Research Announcement for Space Utilization'' promoted by Japan Space Forum.

Reprint requests should be sent to Yasuki Noguchi, Department of Integrative Physiology, National Institute for Physiological Sciences, 38 Nishigonaka, Myodaiji, Okazaki, Aichi, 444-8585, Japan, or via e-mail: noguchi@nips.ac.jp.

REFERENCES

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the ''light-from-above'' prior. *Nature Neuroscience, 7,* 1057–1058.

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology, 14,* 257–262.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception: I. Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9,* 415–430.

Baldo, M. V., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature, 378,* 565–566.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287,* 2036–2038.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8,* 162–169.

Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience, 26,* 4917–4921.

Gregory, R. L. (1997). Knowledge in perception and illusion. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 352,* 1121–1127.

Hamalainen, M., Hari, R., Ilmoniemi, R., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics, 65,* 413–497.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 781–787.

Kanai, R., Sheth, B. R., & Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: Spatial uncertainty is the key to perceptual mislocalization. *Vision Research, 44,* 2605–2619.

Kanai, R., & Verstraten, F. A. (2006). Visual transients reveal the veridical position of a moving object. *Perception, 35,* 453–460.

Kawakami, O., Kaneoke, Y., Maruyama, K., Kakigi, R., Okada, T., Sadato, N., et al. (2002). Visual detection of motion speed in humans: Spatiotemporal analysis by fMRI and MEG. *Human Brain Mapping, 16,* 104–118.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55,* 271–304.

Khurana, B., Carter, R. M., Watanabe, K., & Nijhawan, R. (2006). Flash-lag chimeras: The role of perceived alignment in the composite face effect. *Vision Research, 46,* 2757–2772.

Krekelberg, B., & Lappe, M. (2001). Neuronal latencies and the position of moving objects. *Trends in Neurosciences, 24,* 335–339.

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron, 38,* 317–327.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5,* 910–916.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4,* 432–440.

Mackay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature, 181,* 507–508.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4–18.

Martin, A., & Chao, L. L. (2001). Semantic memory and the brain: Structure and processes. *Current Opinion in Neurobiology, 11,* 194–201.

Mummery, C. J., Patterson, K., Hodges, J. R., & Price, C. J. (1998). Functional neuroanatomy of the semantic system: Divisible by what? *Journal of Cognitive Neuroscience, 10,* 766–777.

Namba, J., & Baldo, V. C. (2004). The modulation of the flash-lag effect by voluntary attention. *Perception, 33,* 621–631.

Nieman, D., Nijhawan, R., Khurana, B., & Shimojo, S. (2006). Cyclopean flash-lag illusion. *Vision Research, 46,* 3909–3914.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370,* 256–257.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences, 6,* 387.

Nishitani, N., & Hari, R. (2002). Viewing lip forms: Cortical dynamics. *Neuron, 36,* 1211–1220.

Noguchi, Y., Inui, K., & Kakigi, R. (2004). Temporal dynamics of neural adaptation effect in the human visual ventral stream. *Journal of Neuroscience, 24,* 6283–6290.

Parviainen, T., Helenius, P., Poskiparta, E., Niemi, P., & Salmelin, R. (2006). Cortical sequence of word perception in beginning readers. *Journal of Neuroscience, 26,* 6052–6061.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396,* 424.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience, 1,* 183–184.

Vandenbulcke, M., Peeters, R., Fannes, K., & Vandenberghe, R. (2006). Knowledge of visual attributes in the right hemisphere. *Nature Neuroscience, 9,* 964–970.

Watanabe, K., Nijhawan, R., Khurana, B., & Shimojo, S. (2001). Perceptual organization of moving stimuli modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 879–894.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1,* 656–657.