Temporal Dynamics of Neural Adaptation Effect in the Human Visual Ventral Stream

Yasuki Noguchi,1,2 Koji Inui,1 and Ryusuke Kakigi1,2
1Department of Integrative Physiology, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan, and 2Department of Physiological Sciences, School of Life Sciences, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0913, Japan

When the same visual stimulus is repeatedly presented with a brief interval, the brain responses to that stimulus are attenuated relative to those at first presentation [neural adaptation (NA)]. Although this effect has been widely observed in various regions of human brain, its temporal dynamics as a neuronal population has been mostly unclear. In the present study, we used a magnetoencephalography (MEG) and conducted a macrolevel investigation of the temporal profiles of the NA occurring in the human visual ventral stream. The combination of MEG with our previous random dot blinking method isolated the neural responses in the higher visual cortex relating to shape perception. We dissociated three dimensions of the NA: activation strength, peak latency, and temporal duration of neural response. The results revealed that visual responses to the repeated compared with novel stimulus showed a significant reduction in both activation strength and peak latency but not in the duration of neural processing. Furthermore, this acceleration of peak latency showed a significant correlation with reaction time of the subjects, whereas no correlation was found between the reaction time and the temporal duration of neural responses. These results indicate that (1) the NA involves the brain response changes in the temporal domain as well as the response attenuation reported previously, and (2) this temporal change is primarily observed as a rapid rising of “what” responses, rather than a temporal shortening of neural response curves within the visual ventral stream as considered previously.

Key words: priming; repetition suppression; visual; ventral pathway; human; magnetoencephalography; MEG

Introduction

One common finding in neurophysiological and neuroimaging studies is a reduced neural response to repeated compared with unrepeated stimuli (Schacter and Buckner, 1998; Wiggs and Martin, 1998; Henson and Rugg, 2003). Although this attenuation was first reported in inferior temporal (IT) neurons of monkeys (Baylis and Rolls, 1987; Brown et al., 1987; Miller et al., 1991; Desimone, 1996; Ringo, 1996), recent studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) showed that the effect also occurs in various regions of the human brain, including the occipital (Grill-Spector et al., 1999; Kourtzi and Kanwisher, 2000), parietal (Naccache and Dehaene, 2001), and inferior frontal cortices (Raichle et al., 1997). In this theory, the neural processing network settles to a stable response more quickly in response to a repeated than novel stimulus, because the network connections involved in producing the response have been reinforced by a previous presentation of the same stimulus. Indeed, a recent fMRI study reported results supporting this view (Henson et al., 2002a), although their analysis requires the assumption of a precise linearity of hemodynamics.

In the present study, we used magnetoencephalography (MEG) with a high temporal resolution and measured directly (BOLD) signal has been regarded as a sign of this effect in fMRI studies, a previous study indicated that this can arise from at least two types of neural activity (Henson and Rugg, 2003) (Fig. 1). One explanation is a reduced firing rate as a whole neural population in each cortical area (Fig. 1a). This view is strongly supported by many studies of unit cell recordings and the “sharpening” theory propounded by Wiggs and Martin (1998). According to this theory, neurons that are not critical for recognizing an object decreases their responses as the object is repeatedly presented, whereas those carrying essential information continue to give a robust response. As a result, the mean firing rate as a whole is attenuated by stimulus repetition. In contrast, a reduction in BOLD responses can also be explained by the response change in the temporal domain: a shortened duration of neural activity (Henson and Rugg, 2003) (Fig. 1b). This account is based on the fact that the hemodynamic response represents the integration of several seconds of neural–synaptic activity and proven to be possible by a previous neural computation theory (Becker et al., 1997). In this theory, the neural processing network settles to a stable response more quickly in response to a repeated than novel stimulus, because the network connections involved in producing the response have been reinforced by a previous presentation of the same stimulus.
the neural responses underlying the neural adaptation (NA) effect in the human shape perception area (Grill-Spector et al., 1999; Kourtzi and Kanwisher, 2001). As a result, our data supported the integrative model of the response reduction and temporal acceleration, as shown in the third hypothesis depicted in Figure 1c.

Materials and Methods

Subjects. Ten healthy volunteers participated in the present study (seven males and three females). All subjects had normal or corrected-to-normal visual acuity. 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. One problem with MEG experiments on visual function is that the neuronal signals from the higher visual areas are difficult to distinguish from those in early visual cortex (such as V1) in most cases because of the insufficient modeling quality of V1 activities. Because neural responses in the early visual areas are relatively insensitive to the repetition effect compared with those in the later visual areas (Buckner et al., 1998; Schacter and Buckner, 1998), the confounding of early visual signals into MEG data in the present study would obscure the NA effect occurring in higher visual areas. Given recent studies reporting that V1 area receives a delayed feedback signal from the higher visual cortex at a latency of 190–230 msec (Noesselt et al., 2002; Halgren et al., 2003) in addition to the primary visual input from the thalamus, it would be difficult to exclude the early visual signals on the basis of signal latency. We therefore presented visual stimuli based on our random dot blinking (RDB) technique developed previously (Okusa et al., 1998). With this method, characters are presented in the center of a black and white random dot field. Although all dots in the field flicker continuously in the resting state, a subset of dots becomes static during the character presentation period, whereas the other dots remain dynamic (Fig. 2a). This static-dynamic contrast enables observers to perceive the shape of a letter. Because the ratio of white and black pixels is fixed throughout both periods, the mean luminance of the field is always the same. Our previous study has shown that this stimulation paradigm effectively inhibits the neural responses from the V1 area and elicits one simple component of magnetic response at a peak latency of ~300 msec, the signal source of which is estimated to lie in the occipito-temporal area around the fusiform gyrus.

We used the RDB method for sequential presentation of two visual stimuli (letters) in the central visual field of the subjects. All stimuli were presented in a random dot field subtending a visual angle of 6 × 6° with a 60 × 60 dots array on the projector screen at a viewing distance of 250 cm. For the dynamic texture, each dot (2 × 2 pixel) changed its position within a 3 × 3 pixel area every 16 msec in a pseudorandom manner to produce vibrating motion. For the static texture, the dots remained stationary. The ratio of white to black pixels was fixed at 1:3 throughout the entire scanning period.

We used six uppercase letters (A, O, E, B, K, P) as letter stimuli. Each letter was used as both S1 (first stimulus) and S2 (second stimulus). The display duration of S1 and S2 was 300 and 500 msec, respectively. The time interval between S1 and S2 (interstimulus interval (ISI)) was either 150, 250, or 350 msec (Fig. 2b), and there were two kinds of trials for each ISI. In SAME trials, the same letter was repeatedly presented as S1 and S2. In DIFF trials, the S1 and S2 letters differed. Because each letter was presented as S1 or S2 in both SAME and DIFF trials at equal times, the difference in brain responses between these two types of trials cannot be attributed to the difference in the visual features of the stimuli presented. Apart from these six conditions (three ISIs for SAME and DIFF), we introduced a control condition in which only S1 was presented for 300 msec (SINGLE condition).

A single scanning session of MEG recordings started with six trials of the SINGLE condition during which subjects were instructed to look passively at the letter presented (no-task period). This period was followed by 72 trials with paired letter stimuli (task period). In this period, stimulus pairs in the six conditions (12 trials for each) were randomly intermixed, and subjects were asked to perform a vowel–consonant...
judgment task with S2, not S1, characters. They were instructed to press one button as quickly as possible when the S2 letter was a vowel (A, O, E) and another button when it was a consonant (B, K, P). All responses were made by the right hand of the subjects. The session ended with another no-task period composed of six trials of the SINGLE condition. To prevent the task and no-task periods from being confused, cue stimuli showing the switch between the two periods were presented. The numerals 2 and 1 were presented at the beginning and end of the task period, respectively. A scanning session containing a total of 84 (50% of the peak). Because it is theoretically independent of variation in the vertical amplitude of the MEG waveform, the FWHM corresponds to an index of the temporal duration of neural activity induced in the ventral visual cortex. Regarding the responses to S1, the peak amplitude and latency were defined as a maximum signal strength (and its latency) in the across-SOI waveforms within the time window of −100 to −500 msec after the S1 onset. The time window was set at 500 to −1200 msec for S2 responses. Once the peak amplitude was determined, we calculated 50% of the peak for each response waveform (half maximum). The time interval between the first and last time points in which signal strength is above the half maximum was defined as FWHM of this waveform.

Apart from the SOI analyses, we estimated the single equivalent current dipole (ECD) on response waveforms to confirm the anatomical location of VEF sources. We adopted a spherical head model based on individual MR images (Hamalainen et al., 1993). The ECDs best explaining the distribution of the magnetic field over at least 20 channels around the signal maxima were searched by the least square method (Wasaka et al., 2002). We accepted only ECDs that accounted for at least 80% of the field variance at the peak.

**Results**

Figure 3a shows the VEF in the SINGLE condition for one subject. Clear deflections were observed primarily in MEG channels on the lateral sides of both hemispheres. Magnetic responses around the occipital pole were relatively small, indicating that...
neural responses in the early visual areas were successfully inhibited by the RDB stimulus. Figure 3b shows the superimposed waveform of all SOIs in the same subject. Consistent with our previous study (Okusa et al., 1998), a large component was observed at a latency of ~300 msec. We plotted in Figure 3c the waveform of all conditions in two representative SOIs marked in Figure 3a. Magnetic responses to S2 that have the same polarity as those to S1 were clearly observed in the six conditions with the paired letter presentation, but they were absent in the SINGLE condition. The S2 response latencies reflected the difference of S2 onset in the three ISI conditions. Within each ISI, the S2 response in the SAME condition appeared to show earlier peak latency (latency at the peak) and smaller deflection than that in the DIFF condition.

The results of dipole analyses indicated that all ECDs calculated on MEG signals in the SINGLE condition were estimated in the vicinity of the occipito-temporal cortex around the fusiform gyrus, which also confirmed our previous results (Okusa et al., 1998). In Figure 4a, the mean dipole location of each hemisphere across subjects was shown on the MR image of a representative subject. According to the head-based coordinate system used, the mean (±SE) coordinates were (−31 ± 9.6, −26 ± 3.5, 44 ± 3.0) for left and (35 ± 3.2, −20 ± 4.2, 48 ± 3.1) for right hemispheres. There was no significant difference of ECD locations between the two hemispheres (p > 0.1, for all axes). These ECD results could be reinforced by two topography maps depicted over the field of 102 sensor positions in MEG. Figure 4b shows a distribution of 222 SOIs selected from the data of 10 subjects and thus represents how many times each sensor was selected as SOI. We also made another contour map (Fig. 4c) in which mean deflection of MEG waveforms at 300 msec of SINGLE condition are plotted for each sensor position. The results of both maps indicate that the signal sources of the 300 msec component are located in the bilateral occipital-temporal regions.

We showed in Figure 5 the across-SOI waveforms of one subject (Fig. 5a) and grand mean of 10 subjects (Fig. 5b). In S2 responses, both peak amplitude and peak latency were clearly different between the SAME and DIFF trials in all three ISI conditions. In contrast, S1 responses in the six conditions with a paired stimulus were almost identical, although, in the grand mean waveform, peak latency was significantly delayed in the SINGLE condition compared with the other six conditions (p < 0.05; t test), probably because of the lack of a task requirement in the SINGLE condition. To closely investigate the temporal profiles of the response waveforms regardless of their amplitude differences, we also presented the grand mean time courses normalized to the S2 peak amplitude of each paired-stimulus condition (Fig. 5c). In all three ISIs conditions, the S2 responses in the SAME trials reach their peak more rapidly than those in the DIFF trials. The SAME responses also precede the DIFF in their signal decreases.

We then examined the NA effect statistically by calculating the three independent parameters on the across-SOI time series of each subject: peak amplitude, peak latency, and FWHM of response waveforms (Fig. 6). In each panel of Figure 6, we used repeated-measures ANOVA of ISI times repetition (SAME vs DIFF) with repetition as a within-subject factor. The Mauchly’s tests indicated that the sphericity assumption was not rejected in all comparisons. In the S1 response, there was no significant main effect or interaction in any parameters (p > 0.05 for all) (Fig. 6a,c,e). In contrast, the peak amplitude and latency of the S2 response showed a significant main effect of repetition (p < 0.0001 for both; the other effects were not significant, p > 0.05) (Fig. 6b,d). However, there was no significant effect in the FWHM of the S2 signal (p > 0.05) (Fig. 6f), although the DIFF trials tended to have a larger FWHM value than the SAME trials (main effect of repetition, F = 4.037; p > 0.05). In Figure 7, we plotted the repetition effect (DIFF-SAME) of the three indices. One-group t test applied to each bar demonstrated that the significant effect of repetition (SAME < DIFF) in S2 peak amplitude and latency could be observed in all three ISIs (Fig. 7a,b) (p < 0.05 for all), whereas no repetition effect of FWHM was observed.
in any ISI conditions (Fig. 7c) (*p > 0.05). The means and SEs in these results were summarized in Table 1.

Although significant repetition effects were observed in S2 peak amplitude and latency, the signal source of these effects remains unclear, because the source location of the repetition effect (DIFF-SAME) was not directly investigated. We therefore conducted the ECD estimation again, using the DIFF-SAME waveform of each ISI condition. The across-subject mean and SE of these ECD locations were: ISI 150 (41 ± 4.5, −24 ± 4.2, 52 ± 4.0), ISI 250 (37 ± 4.8, −15 ± 5.4, 41 ± 3.8), ISI 350 (37 ± 3.8, −21 ± 5.0, 50 ± 4.0) for right hemisphere and ISI 150 (−43 ± 9.7, −20 ± 5.9, 48 ± 5.4), ISI 250 (−27 ± 7.1, −43 ± 7.2, 50 ± 3.8), ISI 350 (−33 ± 8.2, −29 ± 10.3, 49 ± 4.3) for left hemisphere. To examine whether these ECD locations were significantly different from that for the S1 peak, we performed one-way
Table 1. Neural responses to S2 and behavioral RT data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ISI 150</th>
<th>ISI 250</th>
<th>ISI 350</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peak amplitude (fT/cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAME</td>
<td>31.8 (3.9)</td>
<td>33.7 (3.6)</td>
<td>32.5 (2.4)</td>
<td>32.7 (1.9)</td>
</tr>
<tr>
<td>DIFF</td>
<td>39.2 (4.9)</td>
<td>39.9 (3.6)</td>
<td>38.5 (3.2)</td>
<td>39.2 (2.2)</td>
</tr>
<tr>
<td>DIFF-SAME</td>
<td>7.4 (2.1)</td>
<td>6.2 (2.1)</td>
<td>6.0 (1.6)</td>
<td>6.5 (1.1)</td>
</tr>
<tr>
<td>Peak latency (msec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAME</td>
<td>269 (15)</td>
<td>250 (16)</td>
<td>236 (13)</td>
<td>252 (9)</td>
</tr>
<tr>
<td>DIFF</td>
<td>319 (18)</td>
<td>289 (18)</td>
<td>320 (19)</td>
<td>309 (11)</td>
</tr>
<tr>
<td>DIFF-SAME</td>
<td>49 (19)</td>
<td>38 (15)</td>
<td>84 (23)</td>
<td>57 (17)</td>
</tr>
<tr>
<td>FWHM (msec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAME</td>
<td>241 (33)</td>
<td>225 (43)</td>
<td>226 (38)</td>
<td>231 (21)</td>
</tr>
<tr>
<td>DIFF</td>
<td>267 (40)</td>
<td>240 (36)</td>
<td>273 (28)</td>
<td>260 (20)</td>
</tr>
<tr>
<td>DIFF-SAME</td>
<td>26 (27)</td>
<td>15 (21)</td>
<td>47 (27)</td>
<td>29 (14)</td>
</tr>
</tbody>
</table>

ANOVA with four levels (S1 and three ISI conditions) for each ISI condition along with the RT data. The peak latency and RT are measured from S2 onset (same as Fig. 6); mean = SE.

Figure 8. The relationship between behavioral data and MEG waveforms. a, RT data in the vowel–consonant judgment task measured from the S2 onset. As in Figure 6, the blank and filled bars indicate the data in the SAME and DIFF trials, and error bars denote the SE across 10 subjects. **p < 0.0001; paired t test. b, Correlation analyses of the RT data with two temporal parameters of MEG response: S2 peak latency (Fig. 6d) and FWHM (Fig. 6f). The data for six conditions in 10 subjects (60 points) were plotted for each parameter. A significantly high correlation was observed between RT and peak latency (represented as **; r = 0.48; p < 0.0001; significance test for correlation coefficients) but not between RT and FWHM of the S2 response (r = 0.09; p > 0.05).

Discussion

In the present study, we conducted a macrolevel analysis of the temporal profiles of the NA effect in the shape perception area of the human brain. Although brain responses to two rapidly presented visual stimuli were previously indistinguishable from each other because of the limited temporal resolution of other macrolevel imaging techniques such as PET and fMRI, we examined the S1 and S2 responses in the occipito-temporal areas separately by combining MEG recording with the RDB method. In contrast to the S1 response in which no difference was observed between the SAME and DIFF trials, the S2 response showed a clear NA effect both in peak amplitude and peak latency. In contrast, the temporal width (FWHM) of the S2 response was not significantly affected by the stimulus repetition.

One previous study using intracranial event-related potential (ERP) reported that the first distinct response in visual ventral areas appears ~200 msec after the stimulus onset (Puce et al., 1999). In contrast, the occipito-temporal responses in the present study were observed at a latency of 300 msec, 100 msec later than that in the intracranial ERP. We suppose that this delay in response latency would be caused by the special characteristics of our RDB stimuli. Although letters in the ERP study were defined by the luminance difference from the background, our stimuli were presented through a static-dynamic contrast between letters and background. A recent MEG study (Schoenfeld et al., 2003) showed that the response latency in shape processing areas is variable depending on how the visual stimuli are defined. According to them, visual shape stimuli (squares and rectangles) defined by luminance cues activated a serial processing stream from V1 to lateral occipital (LO) and IT regions, whereas shapes defined by motion coherence of random dots (similar but not identical to ours) elicited the activation in MT/V5 before evoking LO and IT responses. As a result, the response latency in LO and IT was delayed for 50–60 msec in response to the motion–defined than luminance-defined shapes, despite the same response latency in early visual areas. Although it is unclear where in the brain our RDB stimuli are processed, it is possible for the RDB letters to be processed in several areas (in the dorsal stream) before reaching occipito-temporal regions, resulting in some delay of activity in the visual ventral pathway.

In Figures 6 and 7, we showed that both peak amplitude and latency in S2 response were significantly modulated by the stimulus repetition. However, these results may be explained merely by the amplitude reduction, because the decrease in peak amplitude inevitably involves the shortening of peak latency when SAME and DIFF show the neural response curves with similar shapes. To examine this possibility, we calculated the correlation coefficient between the peak amplitude and peak latency. If the decrease in peak latency is a “by-product” of the amplitude attenuation, these two values would show a high positive correlation (the smaller the amplitude is, the shorter the latency becomes).

The result revealed that only the weak correlation was observed between peak amplitude and latency (r = 0.17; p > 0.05), suggesting that decrease in amplitude and latency occurred independently, although these changes appear to occur simultaneously in the grand-average time course (Fig. 5b). We also investigated the relationship between DIFF-SAME of peak amplitude and latency, only to find that their correlation was not significant again (r = 0.29; p > 0.05). In addition, the peak amplitude was also poorly correlated with the RT data (r = −0.14; p > 0.05), whereas the high correlation was observed between the peak latency and RT, as shown in Figure 8b. Our conclusion, therefore, is that the...
reductions in peak amplitude and peak latency were independent phenomena produced by the stimulus repetition, and RT decreases behaviorally observed were primarily induced by the decrease in peak latency, not peak amplitude.

One characteristic of the present study is that the temporal durations of neural responses were measured by FWHM. It may seem unnatural that no significant difference between SAME and DIFF was observed in FWHM of S2 responses (Figs. 6, 7), because one would expect a longer processing of DIFF than SAME if the duration of activation was defined by the time interval, during which the amplitudes are above a certain criterion (Fig. 5b). This apparent inconsistency is caused by the fact that the FWHM is determined based on the size of the peak amplitude for each response curve and, thus, is free from the amplitude effects, at least theoretically. Indeed, we initially calculated the temporal duration of waveforms by applying the constant criterion (50% of S1 peak amplitude in the SINGLE condition) to the six conditions with paired stimuli. The result revealed that these measures were highly correlated with the peak amplitudes ($r = 0.62; p < 0.0001$), indicating that temporal and amplitude effects are mixed in this index. In contrast, the correlation between peak amplitude and FWHM is not significant ($r = 0.24; p > 0.05$), as predicted by the theoretical consideration above. Because our primary purpose is to investigate the temporal dynamics of the neural adaptation effect other than response attenuation, we adopted the FWHM as a better index to measure the temporal duration of waveforms regardless of their amplitude difference. However, it may be difficult to argue that the FWHM is completely independent from amplitude effects, because the correlation between them was not zero ($0.24$).

It should be noted that the repetition interval in the present study is relatively short (150–350 msec) compared with those in most previous studies and, therefore, is closely related to the “neural adaptation” technique used in several recent researches (Grill-Spector et al., 1999; Kourtzi and Kanwisher, 2000, 2001). This short ISI also distinguishes our study from previous anatomically constrained MEG studies (Dhond et al., 2001; Marinkovic et al., 2003) using the longer repetition interval that consisted of several intervening items. Indeed, it was suggested that neural response attenuation induced by the immediate stimulus repetition ($< 1$ sec) is qualitatively different from that involved in long-lag repetition of more than several seconds. For example, the neural adaptation with short ISI may be induced by the response attenuation of the same groups of neurons, rather than the decrease in neural population encoding the repeated stimuli as assumed in the sharpening theory (Henson, 2003). Although our study revealed the temporal profiles of the NA effect in visual ventral regions, one should attend to the scale of repetition lag for the interpretation of the present results.

There are two neural models that can account for the BOLD signal decrease in response to the repeated compared with unrepeated stimulus: a reduction in mean firing rate (Wiggs and Martin, 1998) and shortened duration of neural activity (Becker et al., 1997; Henson et al., 2002a). The present results strongly favor the integration of these two models. In addition to a decrease in neural activation by the repeated relative to unrepeated stimulus, we found that NA also induced a temporal shift of peak latency (Fig. 6d). Furthermore, this peak latency showed a significant correlation with RT data (Fig. 8b). These results support the shortened duration model, which attributes NA to the reduced settling time of the neural processing network (Becker et al., 1997). However, our results also showed that this shortening of settling time does not lead to a significant reduction in the temporal duration of response curve itself, because the FWHM, an index of the temporal width of neural processing, was not significantly modulated by the stimulus repetition. These results indicate that, although NA surely involves a change in the neural response in the temporal domain, this change is primarily observed as a speeded response of shape perception areas, rather than a temporal shortening of neural response shapes within these areas.

In conclusion, the present study demonstrated that the NA effect in the human visual ventral stream is characterized by both activation reduction and temporal acceleration, and that the decreased rising time in the ventral visual area is a main factor in the reduction in RT observed. Although there are some limitations on the interpretation of our study (e.g., the frequency changes such as at gamma band were omitted in our trial-averaged analysis, which may relate to hemodynamics changes involved in stimulus repetition), the present results should have an impact on future neural computational theories and provide additional insight into the visual processing mechanism in the human brain.

References


