The development and neural basis of referential gaze perception

Journal Article

The development and neural basis of referential gaze perception

Atsushi Senju, Mark H. Johnson, Gergely Csibra
Centre for Brain & Cognitive Development, Birkbeck, University of London

Correspondence concerning this paper should be sent to: Mark H. Johnson, Centre for Brain and Cognitive Development, School of Psychology, Birkbeck College, Malet Street, London WC1E 7HX. Fax: +44 20 7631 6587; E-mail: mark.johnson@bbk.ac.uk

Abstract

Infants are sensitive to the referential information conveyed by others' eye gaze, which could be one of the developmental foundations of theory of mind. To investigate the neural correlates of gaze-object relations, we recorded ERPs from adults and 9-month-old infants while they watched scenes containing gaze shifts either towards or away from the location of a preceding object. In adults, object-incongruent gaze shifts elicited enhanced ERP amplitudes over the occipito-temporal area (N330). In infants, a similar posterior ERP component (N290) was greater for object-incongruent gaze shifts, which suggests that by the age of 9 months infants encode referential information of gaze in a similar way to adults. In addition, in infants we observed an early frontal ERP component (anterior N200), which showed higher amplitude in response to the perception of object-congruent gaze shifts. This component may reflect fast-track processing of socially relevant information, such as the detection of communicative or informative situations, and could form a developmental foundation for attention sharing, social learning and theory of mind.

Acknowledgment

Introduction

Human eyes have a unique morphology compared to other primate species. The sclera is white (i.e. it lacks pigmentation) and widely exposed, which makes a high contrast with the coloured iris (Kobayashi & Kohshima, 1997, 2001). As a result, it is very easy to infer gaze direction from human eyes. Considering this unique morphology, it is not surprising that eye gaze plays a crucial role in human social interaction and communication (Baron-Cohen, 1995;

Csibra & Gergely, 2006; Kleinke, 1986). Humans are sensitive to where another is looking (Gibson & Pick, 1963), shift attention to the same direction as another's gaze (Driver, Davis, Ricciardelli, Kidd, Maxwell & Baron-Cohen, 1999; Friesen & Kingstone, 1998; Langton & Bruce, 1999) and rely heavily on the information from others' eye region when inferring their mental states (Baron-Cohen, 1995; Baron-Cohen, Wheelwright & Jolliffe, 1997). Moreover, atypical processing of others' eye gaze has been reported in the population with autism spectrum disorders (Baron-Cohen, 1995; Baron-Cohen, Wheelwright et al., 1997; Senju, Hasegawa & Tojo, 2005; Senju, Tojo, Dairoku & Hasegawa, 2004; Senju, Yaguchi, Tojo & Hasegawa, 2003), who suffer from severe difficulties in social interaction and communication. These findings suggest a strong link between eye gaze processing and social-communicative development.

One of the most important aspects of eye gaze is its referential nature. Eye gaze points to the target of another's attention and intention, and help us identify and share attention and communicative topics (Baron-Cohen, 1995; Bloom, 2000). Understanding the referential nature of gaze may also be a requirement for the development of theory of mind. For example, Baron-Cohen (1994, 1995) argues that the shared attention mechanism (SAM), which represents triadic relations between self, other and the object, is the basis of the development of the theory of mind mechanism (ToMM), which infers others' mental states. Recent neuroimaging studies revealed that the superior temporal sulcus (STS), as well as the inferior parietal sulcus (IPS) and the fusiform gyrus (FFG), are involved in the encoding of the relations between gaze direction and the location of the object being looked at (Pelphrey, Singerman, Allison & McCarthy, 2003). These regions showed greater activity in response to gaze shifts which were incongruent, compared to congruent, with the location of randomly appearing objects. The involvement of STS in gaze processing has also been reported elsewhere (Hoffman, & Haxby, 2000; Pelphrey, Viola, & McCarthy, 2004; Puce, Allison, Bentin, Gore, & McCarthy, 1998; Wicker, Michel, Henaff, & Decety, 1998). However, little is known about the temporal course of gaze processing in the brain, end especially about the encoding of gaze-object relations.

The developmental course of the processing of gaze-object relations is another important question. Although the gaze-object congruency effect was replicated in 7- to 10-year-old children (Mosconi, Mack, McCarthy & Pelphrey, 2005), no studies have examined the neural correlates of the processing of gaze-object relations below that age. Thus, it remains unknown whether younger children, or even infants, can encode gaze-object relations. In typical development, sensitivity to the eyes of other humans appears from very early on. Even newborns prefer direct over averted gaze (Farroni, Csibra, Simion & Johnson, 2002) and show a rudimentary form of gaze following (Farroni, Massaccesi, Pividori & Johnson, 2004). By at least 4 months, infants also show enhanced cortical processing of faces accompanied by direct gaze (Farroni et al., 2002; Farroni, Johnson & Csibra, 2004). Finally, a recent behavioural study (Senju, Csibra & Johnson, 2006) revealed that 9-month-old infants are sensitive to gaze-object relations. In the present study we investigated whether 9-month-old infants show differential neural processing of object-congruent and object-incongruent gaze shifts.

The aim of the current study is to investigate the neural correlates of the processing of gaze-object relations and its time-course both in adults and in infants with event-related potential (ERP) measurement. There are two main reasons to choose ERP measurement in this study. First, ERPs have a good temporal resolution, which enables us to examine the detailed time-course of the gaze processing. Secondly, ERPs are the most feasible way to measure brain activity from awake human infants (de Haan, Johnson & Halit, 2003; Johnson, de Haan, Oliver, Smith,

Hatzakis, Tucker & Csibra, 2001). The aim of Experiment 1 was to extend the study of Pelphrey et al. (2003) to ERP measurement. Previous ERP studies focused on the perception of gaze shifts (Puce, Smith & Allison, 2000) or other types of biological motion (Hirai, Senju, Fukushima & Hiraki, 2005; Jokisch, Daum, Suchan & Troje, 2005) consistently found occipito-temporal negativities around 200 and 300 ms which were assumed to reflect neural activity in the STS. In the present study, we predicted that perceived gaze shifts would elicit an occipito-temporal negative component, which would be of greater amplitude in response to the object-incongruent, compared to object-congruent, gaze shifts. In Experiment 2, 9-month-old infants were tested with similar stimuli as in Experiment 1. Since previous ERP studies found that posterior components between 200 and 300 ms in young infants are sensitive to gaze direction (Farroni et al., 2002, 2004a) and biological motion (Hirai & Hiraki, 2005; Reid, Hoehl & Striano, 2006), we predicted that perceived gaze shifts would elicit ERP components over the posterior electrodes and that these components would discriminate object-congruent from object-incongruent gaze shifts.

Experiment 1

The aim of Experiment 1 was to explore the effects of gaze-object congruency on ERPs. Our stimuli were similar to those used in Pelphrey et al. (2003): An image of a female face was presented on the centre of the screen, and an object occasionally appeared to the periphery. Then, the face shifted her gaze either toward the object, or away from it. Moreover, several modifications were made from Pelphrey, Morris and McCarthy (2005). First, in our experiment, the object disappeared from the screen at the same time when the gaze shift began in order to avoid confounds from enhanced object processing due to the gaze cueing effect (Reid, Striano, Kaufman & Johnson, 2004; Schuller & Rossion, 2001). If the object had been left on after the gaze shift, object-congruent gaze would have triggered attention to the object and facilitate object processing, which would have resulted in differential ERPs between congruent and incongruent trials without necessarily requiring referential understanding. Note that the short temporal gap (240 ms) between the appearance of the object and the gaze shift ensured that gaze shift would be interpreted as a response to, and hence referring to, the object. In the second modification, we added two control conditions (object only condition, gaze only condition) to study the effect of object onset and gaze shift on ERPs. Comparison between the experimental and control conditions would discriminate the effect of gaze-object congruency from the summation of object- and gaze-related ERPs. Finally, we analyzed ERPs for leftward and rightward gaze shifts separately, since both object-related (Eimer, 1994; Schuller & Rossion, 2001) and gaze-related (Puce et al., 2000) ERPs are lateralized according to the visual field in which stimuli are presented.

Methods
Participants
Twenty-one healthy right-handed volunteers participated in this study. Sixteen of them (10 female, 6 male; age range: 19 - 41 years old; average age: 28.5 years) provided an adequate number of artifact-free ERP trials and were included in the analysis. All participants had normal or corrected to normal visual acuity. Participants provided written informed consent, and were paid for their participation.
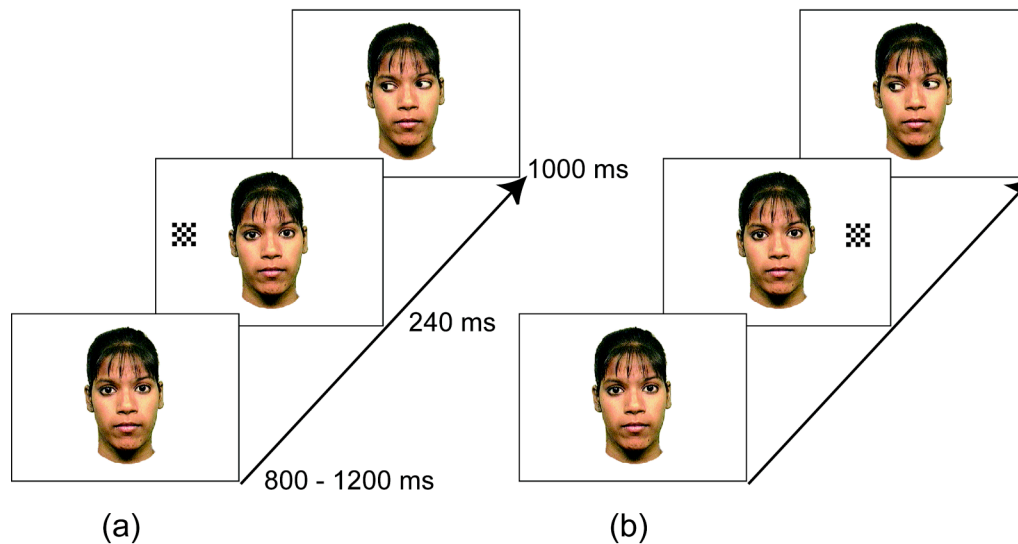
Figure 1. Examples of stimulus sequences of (a) 'gaze-congruent' and (b) 'gaze-incongruent' trials in Experiment 1.

Stimuli and Procedure

Colour photographs of a female face (7º wide and 12º high, with each eye subtending 1.5º wide 0.7º high) were used as the stimuli. Each stimulus sequence started with a face with direct gaze, which was presented for a varying duration between 800 and 1200 ms. There were five types of trials: two experimental conditions (Figure 1). where both gaze shifts and objects were presented on either congruent or incongruent sides, two control conditions. where only a gaze shift or an object was presented, and a target condition. Each trial started with a face with direct gaze, which was presented for a varying duration between 800 and 1200 ms. In the experimental conditions, a black and white checkerboard, which subtended to 2º, then appeared either to the right or to the left of the face at 10º away from the centre. The checkerboard disappeared after 240 ms, and at the same time the eyes of the face shifted toward the location where the object was presented (congruent condition) or toward the opposite direction (incongruent condition) and remained in that position for 1000 ms. Control conditions were the same as the experimental conditions except that the eyes did not move (object only condition) or the checkerboard did not appear (gaze only condition). The stimulus sequence in the target condition was the same as in test conditions except that the eyes closed, rather than shifted laterally, 240 ms after the onset of the checkerboard. Participants were asked to fixate on the face and press a response pad whenever they detected the target event (i.e. the eyes closing). The purpose of the target condition was to keep the participants' attention on the face. Neither behavioural nor electrophysiological responses to the target trials were analyzed.

The EEG recording consisted of eight blocks. Within each block, 30 trials of each condition were presented in a randomized order, with location of checkerboard presentation and/or the direction of eye movement counterbalanced.

EEG recording and analysis

Continuous electroencephalogram (EEG) and electro-oculogram (EOG) were recorded using a Geodesic Sensor Net (GSN; Electric Geodesics, Eugene, OR) with 128 electrodes. All

the recordings were initially referenced to the vertex. The signal was amplified by EGI NetAmps amplifier (Electric Geodesics, Eugene, OR) with a bandpass filter of 0.1 - 100 Hz, digitized at a 250 Hz sampling rate, and stored in computer disk for off-line analyses.

The continuous EEG was segmented from 400 ms before the onset of gaze shift (i.e., 160 ms before the onset of checkerboard presentation) to 900 ms after the onset of gaze shift. Data from each EEG electrode was discarded when signal variation exceeded 50 µV. Whole trials were discarded when more than 10 EEG electrodes were discarded or signal variation in EOG electrodes exceeded 50 µV. Otherwise, missing data were interpolated using spherical spline interpolation. (Data were also analyzed with missing channels excluded rather than interpolated, but this modification did not change the results.) Data were baseline-corrected against the average voltage of 160 ms prior to the onset of checkerboard (i.e. from 400 ms to 240 ms prior to the onset of gaze shift). Each segment was averaged for each condition for each participant and re-referenced to the average of the signal of all electrodes. Note that ERPs were derived separately for left/right of the gaze shift and/or the location of the checkerboard.

Grand-averaged data, as well as individuals' average ERPs, was visually inspected and three time windows were chosen to capture the components sensitive to eye-object congruency: P120 (100 - 150 ms) in midline lower-occipital region (electrode number 75, 82 and 83), posterior N200 (170 - 230 ms) and N330 (310 - 350 ms) in lateral (left: 56, 57, 63; right: 100, 101, 108) and semi-medial (left: 58, 59, 64, 65, 69, 70; right: 90, 91, 92, 95, 96, 97) posterior regions. The peak amplitudes and latencies of each electrode were detected within the specified time windows. Then, within the selected channel groups, the electrode with the largest peak amplitude was selected to represent each component.

In order to further test the effect of eye-object congruency, difference waves were calculated by subtracting the corresponding control conditions from the experimental conditions. For example, from the ERPs of incongruent condition with leftward gaze shift, both the ERPs of gaze only condition with leftward gaze shift and that of object only condition with rightward object presentation were subtracted. Peak amplitudes and latencies of each component of difference waves were derived the same way and subjected to the same ANOVAs as the ERPs.

(a) ERPs for leftward gaze shift

——— object-congruent gaze shift
——— object-incongruent gaze shift

(b) ERPs for rightward gaze shift

Figure 2. Grand-average ERP waveforms for (a) leftward gaze shift and (b) rightward gaze shift at channels corresponding with 10-20 sites.

(a) ERPs for leftward gaze shift

(b) ERPs for rightward gaze shift

∇ : P120
▲ : N200
△ : N330

object-congruent gaze shift
object-incongruent gaze shift

Figure 3. Grand-average ERP waveforms for (a) leftward gaze shift and (b) rightward gaze shift at channel groups included in the analyses in Experiment 1. In addition to the waveforms at channel groups included in the analyses, ERPs at Fz, Cz and Pz are presented to illustrate the spatial distributions of the ERPs.

(a) Difference waves for leftward gaze shift

object-congruent gaze shift
object-incongruent gaze shift

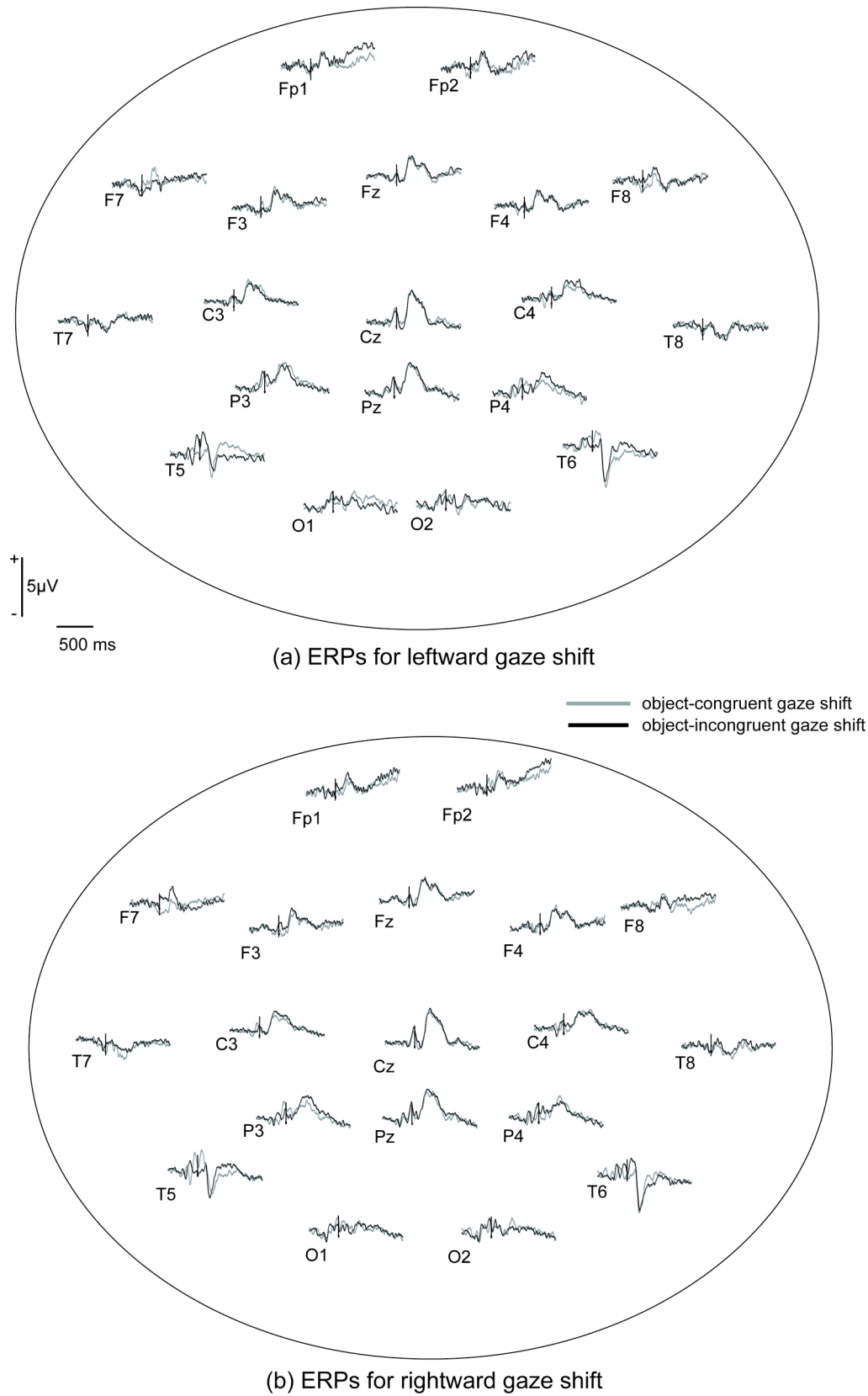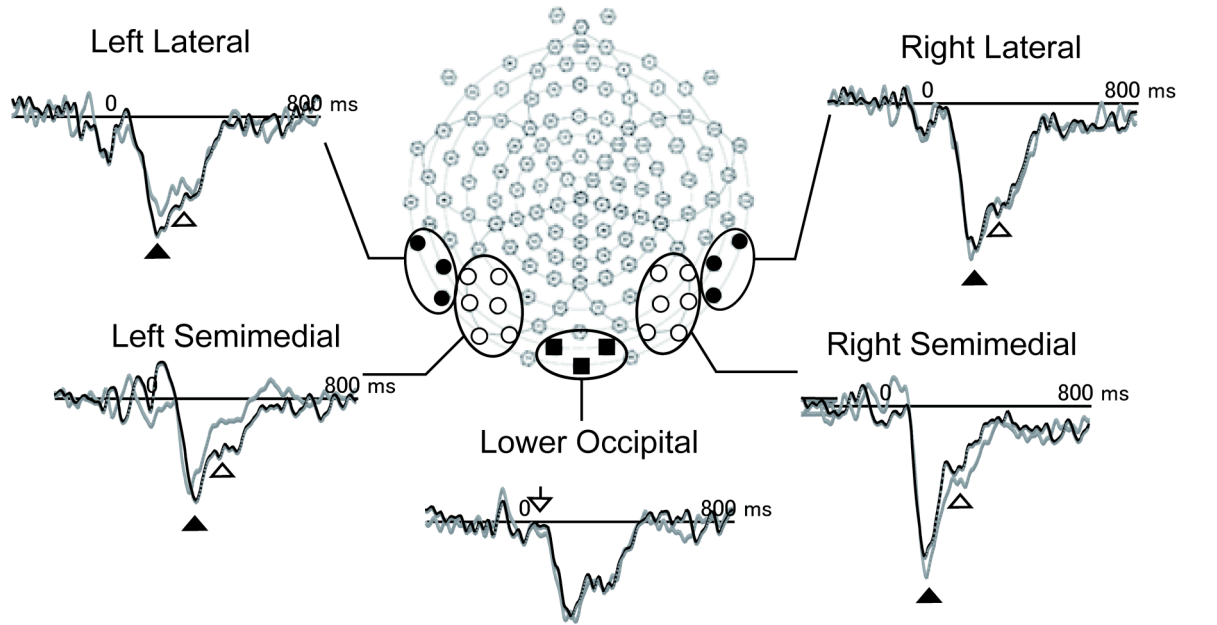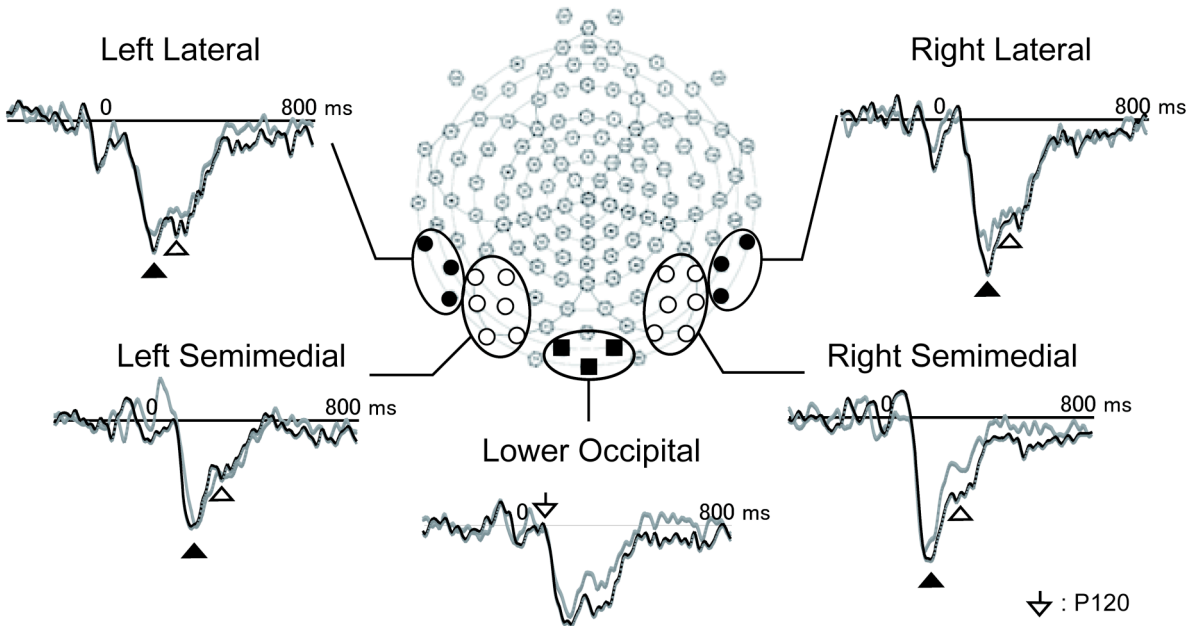(b) Difference waves for rightward gaze shift

Figure 4. Grand-average difference waveforms for (a) leftward gaze shift and (b) rightward gaze shift at channels corresponding with 10-20 sites.

(a) Difference waves for leftward gaze shift

(b) Difference waves for rightward gaze shift

∀ : P120d
▲ : N200d
△ : N330d

2μV
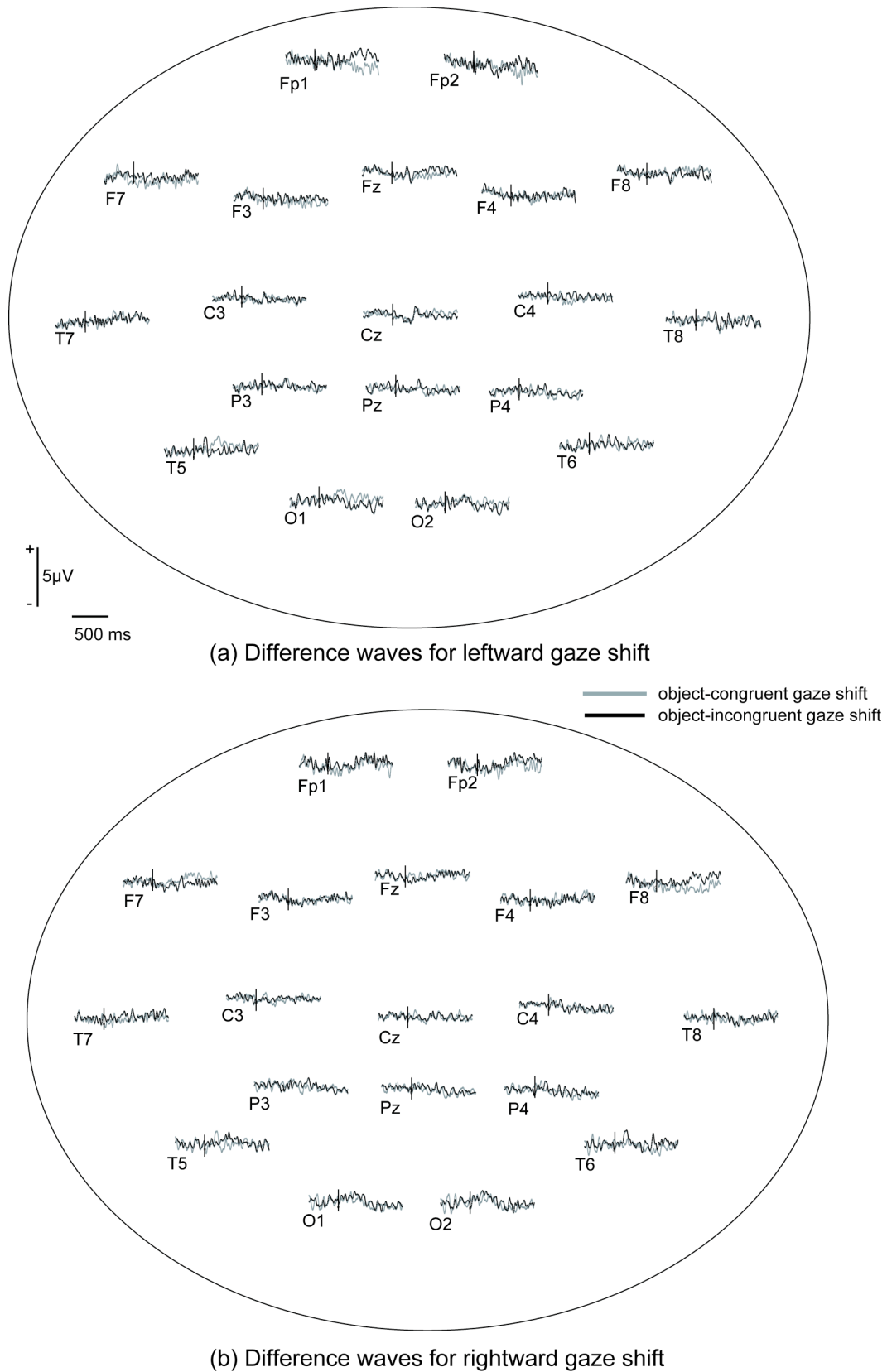
object-congruent gaze shift
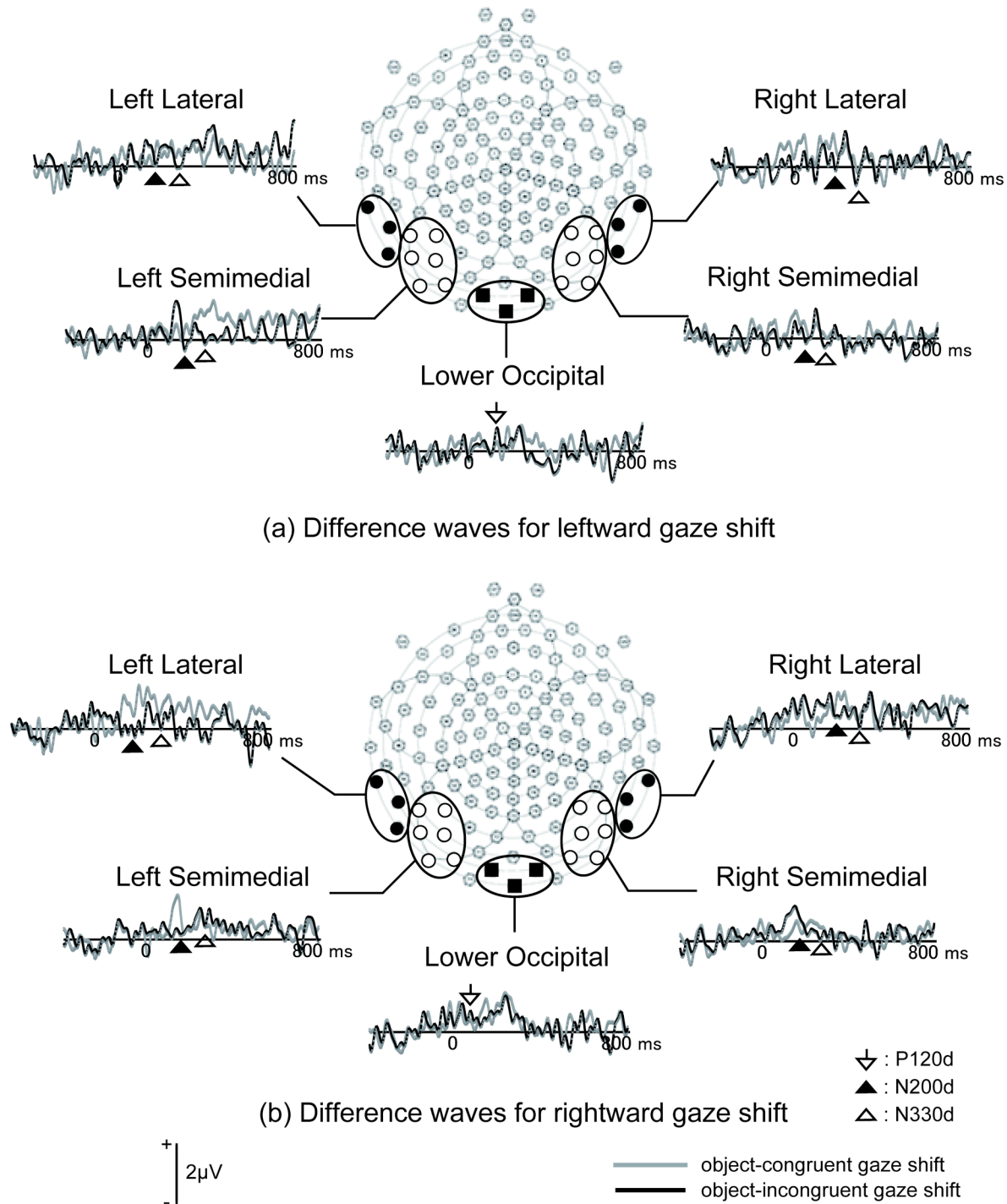object-incongruent gaze shift

Figure 5. Grand-average difference waveforms for (a) leftward gaze shift and (b) rightward gaze shift at channel groups included in the analyses in Experiment 1.

Results

ERPs

Figure 2 illustrates the spatial distributions of the grand average ERPs and Figure 3

depicts grand average ERPs at selected channel groups. First, the peak amplitudes and latencies of the midline occipital P120 in the test conditions were subjected to two-way ANOVAs with congruency (congruent or incongruent) and gaze direction of the stimuli (left or right) as within-subject factors. According to this analysis, the incongruent gaze shift elicited larger P120 amplitudes than did the congruent gaze shift ($F(1,15) = 6.74$, $p < .05$, $\eta_p^2 = .310$). For the peak amplitudes and latencies of P120, no other main effects or interactions were significant.

The peak amplitudes and latencies of the posterior N200 and N330 in the experimental conditions were subjected to four-way ANOVAs with congruency (congruent or incongruent), gaze direction (left or right), hemisphere (left or right) and laterality (semi-medial or lateral) as within-subject factors. Importantly, neither peak amplitudes nor latencies of N200 showed significant effects related to congruency. The N200 was larger ($F(1,15) = 11.77$, $p < .01$, $\eta_p^2 = .440$) and slower ($F(1,15) = 30.44$, $p < .01$, $\eta_p^2 = .670$) in semi-medial channels than in lateral channels. In addition, an interaction between gaze direction and hemisphere was significant for peak latencies ($F(1,15) = 6.66$, $p < .05$, $\eta_p^2 = .307$), because the posterior N200 has a shorter latency in the hemisphere contralateral to the direction of the gaze shift.

The peak amplitudes of N330 showed a significant main effect of congruency ($F(1,15) = 5.35$, $p < .05$, $\eta_p^2 = .263$) and a three-way interaction between congruency, gaze direction and hemisphere ($F(1,15) = 14.83$, $p < .01$, $\eta_p^2 = .497$). Simple effect analyses revealed that the main effect of congruency was significant in the hemisphere ipsilateral to the gaze direction ($p < .05$) but not in the contralateral hemisphere ($p > .1$). No other main effect or interaction, nor effects for the peak latencies of N330, were significant.

Difference waves

Figure 4 illustrates the spatial distributions of the grand averages of difference waves and Figure 5 depicts grand averages for difference waves at selected channel groups. Again, incongruent gaze shifts elicited larger P120d than did congruent gaze shifts ($F(1,15) = 6.36$, $p < .05$, $\eta_p^2 = .298$). For the peak amplitudes and latencies of P120d, no other main effects or interactions were significant.

For peak amplitudes of the posterior N200 in difference waves (N200d), the main effect of congruency ($F(1,15) = 9.51$, $p < .01$, $\eta_p^2 = .239$) and congruency by gaze direction interaction ($F(1,15) = 6.74$, $p < .05$, $\eta_p^2 = .388$) were significant. This interaction reflects the fact that incongruent gaze shifts elicited larger N200d than the congruent gaze shifts in response to leftward gaze shifts ($p < .05$) but not in response to rightward gaze shifts ($p > .1$). No other main effect or interactions were significant. For the peak latencies of N200d, a three-way interaction between congruency, hemisphere and laterality was significant ($F(1,15) = 7.00$, $p < .05$, $\eta_p^2 = .318$). However, the simple main effect of congruency was not significant in any locations. In addition, the main effect of hemisphere ($F(1,15) = 5.11$, $p < .05$, $\eta_p^2 = .254$) and the gaze direction by laterality interaction ($F(1,15) = 7.80$, $p < .05$, $\eta_p^2 = .342$) were also significant.

As with the N330, the N330 in difference waves (N330d) was larger for incongruent gaze shifts than for congruent gaze shifts ($F(1,15) = 17.68$, $p < .01$, $\eta_p^2 = .541$). The interaction between congruency and gaze direction ($F(1,15) = 20.00$, $p < .01$, $\eta_p^2 = .571$) and the three-way interaction between congruency, gaze direction and hemisphere ($F(1,15) = 10.36$, $p < .01$, $\eta_p^2 = .409$) were also significant. Simple effect analyses revealed that the N330d for incongruent gaze shifts was larger than that for congruent gaze shifts in response to leftward gaze shifts (Left hemisphere: $p < .001$, Right hemisphere: $p < .05$) but not in response to the rightward gaze shifts

($p > .1$). The N330d latencies were not affected by congruency, as no main effects or interactions were significant except for the main effect of laterality (F (1,15) = 5.24, $p < .05$, $\eta_p^2 = .259$) in this analysis.

Discussion

      Object-incongruent gaze shifts elicited larger amplitudes at the P120 and N330 compared to the object-congruent condition. This effect is unlikely to have been due to mere summation of ERPs related to the peripheral object and following gaze shift to the contralateral direction, because the same effect was found in the difference waves (in which ERPs in the object-only and gaze-only conditions were subtracted from ERPs in experimental conditions). Rather, the P120 and the N330 reflect the encoding of the relations between the object location and gaze direction. Another interesting finding is that in the difference waves, but not in original ERPs, the peak amplitude of the posterior N200 was larger for the object-incongruent gaze shift than for the object-congruent gaze shift. These results extend previous fMRI studies (Pelphrey et al., 2003, 2005) and demonstrated that gaze-object congruency modulates ERPs in adults.

      With regard to the cerebral source of these effects, based on the previous fMRI studies with similar stimuli (Pelphrey et al., 2003, 2005), one possibility is that the N330 mainly reflects activity in STS. Previous ERP studies with other types of biological motion (e.g. point-light walker) also found an occipito-temporal negativity around 300 ms after the stimulus onset (N330: Hirai et al., 2005; N300: Jokisch et al., 2005), suggesting that the present N330 relates to the activity in the STS. Moreover, the STS seems to encode the relations between perceived biological motion and environmental context (Pelphrey, Morris & McCarthy, 2004; Saxe, Xiao, Kovacs, Perrett & Kanwisher, 2004), which also suggests that modulation of the N330 can reflect the encoding of a gaze shift in relation to the surrounding context (i.e. location of preceding object), in STS. In addition to the STS, the IPS (Pelphrey et al., 2003, 2005) and the FFG (Pelphrey et al., 2003) have been reported to show larger activation in response to object-incongruent gaze shifts, which could also contribute to the present effects.

      In the N330d, when both of the gaze-specific component and object-specific component were subtracted, only the leftward gaze, and not the rightward gaze, showed an effect of gaze-object congruency. This observation could be related to the left visual field dominance of gaze perception (Ricciardelli, Ro & Driver, 2002). Although we did not find a significant right-hemisphere dominance either in N330 or in N330d, previous fMRI studies found greater right hemisphere activity in gaze processing (Pelphrey et al., 2003, 2005). We speculate that hemispheric asymmetry affected an earlier stage of visual processing, which modulated the gaze-object congruency effect in N330d. Further studies will be required to examine the time course of hemispheric lateralization of the gaze processing.

      The modulation of the P120 and the posterior N200 seem to be related to the perception of motion onset. In previous studies, perception of motion onset evoked positive deflection (P1) around 100-130 ms and a subsequent negative deflection (N2) around 150-200 ms (Bach & Ullrich, 1994, 1997). The former component relates to early visual processing in V1/V2 and the latter may reflect motion-specific processing. It is interesting that the P120 amplitudes were modulated by gaze-object congruency. This effect cannot be attributed to the combination of the laterality effect of object location and gaze direction because the effect was still significant in difference waves. Related effects have been reported previously. Klucharev and Sams (2004) reported that posterior positive deflection at 110 ms is larger in ERPs to straight than in ERPs to

right averted gaze, which is consistent with our finding in that P1 is sensitive to perceived gaze direction. Several other studies found that the P1 is sensitive to various kinds of social information. Its amplitudes are larger for inverted than for upright face (Itier & Taylor, 2004; Linkenkaer-Hansen, Palva, Sams, Hietanen, Aronen & Ilmoniemi, 1998) and larger for expressive than for neutral face (Batty & Taylor, 2003). In addition, Meeren, van Heijnsbergen and de Gelder (2005) found that incongruence between facial and bodily expressions elicits larger P1 amplitudes, which may corroborate the present study and suggests that the P120 can reflect an early contextual modulation of social information processing.

Experiment 2

In Experiment 1 we recorded ERPs from adult participants while they were viewing gaze shifts either toward or away from the location of a preceding object. Our results extended those of Pelphrey et al. (2003, 2005) and revealed that object-incongruent gaze shifts elicit larger amplitudes in some posterior ERP components. In Experiment 2, we aim to extend this finding and investigate whether infants show similar neural activations in response to perceived gaze shifts. Stimuli were identical to those used in Experiment 1, but only the experimental conditions were presented because of the shorter attention span of infants compared to adults. We predicted that gaze-object congruency would affect ERPs in 9-month-old infants similarly to that of adults because our previous behavioural study found that infants at this age can discriminate the scenes with object-congruent gaze shifts from those with object-incongruent gaze shifts (Senju et al., 2006).

Methods
Participants
Ten 9-month-old infants (4 female, 6 male; age range: 260 - 282 days old; average age: 274 days) provided an adequate number of artifact-free ERP data and were included in the analysis. An additional 23 infants also participated the study but were excluded from the analyses due to fussiness (6), large eye and/or body movements that resulted in the recording artifacts (12), intolerance of the EEG net (3) or experimenter error (2).

Stimuli and Procedure
Stimuli used for the recording were the same as those used in Experiment 1, except that colour pictures of fish (2º by 3.8º on the screen), rather than black and white checkerboards, were presented as objects, and their appearance was accompanied by a brief sound delivered from behind the centre of the screen. Only the two experimental conditions (congruent and incongruent) were presented. The location of the object presentation was randomized across trials.

Infants sat on the caretaker's lap, approximately 60 cm from the screen, and passively viewed the stimuli. At the beginning of the recording session, the infants' attention was drawn to the same colour picture of a female face as the one used in recording, with eyes blinking every 400 ms, and alerting sounds delivered from behind the screen. An experimenter monitored the infant's face through a video camera, which was located beneath the screen throughout the recording. The experimenter pressed a key and presented the stimuli only when the infant was

attending to the screen. When the infant looked away from the screen, the experimenter stopped the stimulus presentation and attempted to recapture the infant's attention by presenting the face with blinking eyes sounds. The experiment was terminated when the infants became too fussy. The average numbers of trials presented was 203 (range: 145-266).

EEG recording and analyses

EEG and EOG were recorded exactly the same way as Experiment 1 except that a 62-channel sensor net, instead of 128-channel net, was used. As in Experiment 1, the continuous EEG was segmented from 400 ms before the onset of gaze shift (i.e. 160 ms before the appearance of the fish) to 900 ms after the onset of gaze shift. Data from each EEG electrode was manually edited for electrical and movement artifact. Whole trials were discarded when more than 6 EEG electrodes were discarded or when the signal contained eye movements. Otherwise, missing data were interpolated using spherical spline interpolation. In addition, video recordings of each infant's behaviour were coded off-line to ensure that they were fixating on the screen. Data were baseline-corrected against the average voltage of 160 ms prior to the onset of the object (i.e. from 400 ms to 240 ms prior to the onset of gaze shift). Each segment was averaged for each condition for each participant and re-referenced to the average of all electrodes. Due to the smaller numbers of trials, trials with both leftward and rightward gaze shift were averaged together to obtain reliable ERPs in infants. The mean numbers of trials included for averaging were 36 for congruent trials (range: 11-92) and 37 for incongruent trials (range: 11-89).

The grand-averaged ERPs, as well as individual ERPs, were visually inspected and three time windows were chosen to capture the components sensitive to eye-object congruency: anterior N200 (150 - 250 ms) and N400 (350 - 450 ms) in the anterior regions (frontal: 6, 7, 10, 11; front-central: 3, 4, 8, 9, 58), and N290 (100 - 300 ms) and P400 (350 - 450 ms) in the posterior regions (left: 18, 22, 25, 28, 29, 32, 33, 36,37; midline: 30, 34, 38; right: 40, 41, 42, 43, 44, 45, 46, 47, 50). As in Experiment 1, each channel grouping was assigned based on the visual inspection of grand-average, as well as individual, ERP waveforms. Since each peak was more broadly and equally distributed in infants than in adults, and since fewer trials were averaged in infants than in adults, peak amplitudes and latencies were determined for each electrode within specified time windows and averaged within selected channel groups to increase the signal-to-noise ratio.

Results

Figure 6 depicts the grand average ERPs obtained in this study. Peak amplitudes and latencies of each component were subjected to two-way ANOVAs with congruency (congruent or incongruent) and channel group (frontal or front-central for anterior components; left, centre or right for posterior components).

The main effect of congruency was significant for the peak amplitudes of the anterior N200 (F (1,9) = 19.26, $p < .01$, $\eta_p^2 = .681$), N290 (F (1,9) = 7.62, $p < .05$, $\eta_p^2 = .458$) and N400 (F (1,9) = 8.55, $p < .05$, $\eta_p^2 = .487$), and was marginally significant for P400 amplitude (F (1,9) = 5.09, $p = .051$, $\eta_p^2 = .361$). The anterior N200 and N400 amplitudes were larger for congruent gaze shifts, but the N290 and P400 were larger for incongruent gaze shifts. Non-parametric Wilcoxon's Sign Rank Tests confirmed these results (N200: $Z = -2.80$, $p < .01$; N290: $Z = -1.99$, $p < .05$; N400: $Z = -8.55$, $p < .05$; P400: $Z = -1.89$, $p = .059$). Neither the main effects of channel group, nor congruency by channel group interactions, reached significance in any analyses. No

main effects or interactions reached significance in the peak latencies of any component.
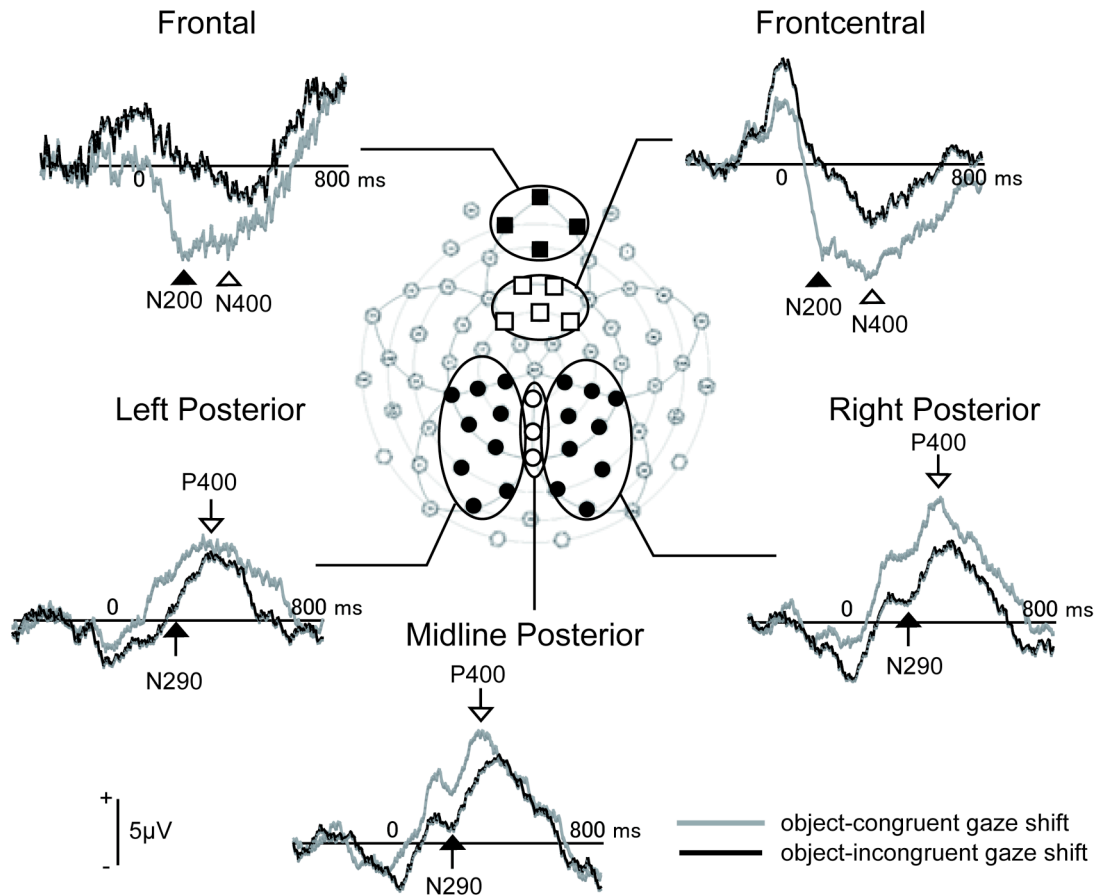

Figure 6. Grand-average ERP waveforms at selected channel groups in Experiment 2.

Discussion

These results clearly demonstrate that gaze-object congruency affects neural processing of perceived gaze shifts in 9-month-old infants. This finding supports our previous behavioural study (Senju et al., 2006) and we conclude that infants can encode the relation between the location of an object and the direction of a gaze shift by at least 9 months of age. Three main ERP components discriminated gaze-object congruency. Posterior N290 amplitudes were greater for object-incongruent gaze shifts compared to object-congruent gaze shifts, a result similar to that seen with the adult N330. The anterior N200 and N400 components, in contrast, were larger for object-congruent compared to object-incongruent gaze shifts. No such anterior components were found sensitive to the gaze-object congruency in adult ERPs.

The response properties, latency and scalp distribution of the N290 suggest that it may be the precursor of the adult N170 (de Haan, Pascalis & Johnson, 2002; Farroni et al., 2002, 2004a; Halit, Csibra, Volein & Johnson, 2004). N290 increase has been found in response to the onset of face presentation and is modulated by the gaze direction of faces (Farroni et al., 2002, 2004a). Halit et al. (2004) also reported that the N290 is a more sensitive measure of face processing than the P400. It is thus possible that the N290 reflects the encoding of gaze direction in relation to

the location of the preceding object.

The greater amplitude of negative components in response to object-congruent gaze shift was found in anterior scalp regions. These effects were long-lasting (the effect was found at both 200 ms and 400 ms after the onset of the gaze shift). It is difficult to interpret the function of these components since this is the first study investigating ERPs to gaze shift perception in infants. We did not find a corresponding frontal component in adults, but previous studies may suggest that these components reflect face/gaze processing. Johnson, Griffin, Csibra, Halit, Farroni, de Haan, Tucker, Baron-Cohen and Richards (2005) conducted source separation and localization analyses on the data set of infants viewing faces and found at least one prefrontal source which discriminated between upright faces with direct gaze from those with averted gaze. Moreover, the peak latency of this prefrontal component was around 200 ms in 12-month-old infants (Richards, personal communication). This putative prefrontal source appears similar to the current N200 in latency, topography and sensitivity to the gaze direction. Webb, Long and Nelson (2005) also reported early ERP component (Pb: 150-350 ms for 8- to 12-month-old infants) sensitive to facial familiarity, which distributed widely across the scalp. In addition, prefrontal activity in response to face perception has also been reported in another neuroimaging study (Tzourio-Mazoyer, De Schonen, Crivello, Reutter, Aujard & Mazoyer, 2002). Such early prefrontal activities could be the neural correlates of subcortical face processing involving amygdala (Johnson, 2005). However, this is a preliminary finding and further research will be required to examine the functions and generators of this early component recorded over frontal sites.

General Discussion

The present study is the first to record ERPs in response to perceived gaze-object relations in adults and infants, and revealed that both adults and infants encode the congruence between gaze direction and the location of an object. The adult results extended those of preceding fMRI studies (Pelphrey et al., 2003, 2005) by showing that an occipito-temporal component (N330), which may reflect activity from STS, is increased in response to object-incongruent, as compared to object-congruent, gaze shifts. Moreover, a similar posterior negativity (N290) was recorded in 9-month-old infants, and amplitudes were also greater for object-incongruent gaze shifts. Thus, the current study extended previous findings that the infant N290 is modulated by the gaze direction (Farroni et al., 2002, 2004a) and is the first to reveal that the N290 discriminates gaze-object relations. The results also suggest that infants, as adults, recruit a posterior cortical network to encode referential information of gaze by the age of 9 months. Since the referential understanding of others' gaze is essential for sharing attention, the neural mechanisms encoding referential information from gaze may be a foundation for the development of SAM (Shared Attention Mechanism) and ToMM (Theory of Mind Mechanism) (Baron-Cohen, 1994, 1995).

In both adults and infants, the amplitudes of some posterior negativities were larger for object-incongruent than for object-congruent gaze shifts. This is consistent with the findings of Pelphrey and colleagues (Pelphrey et al., 2003, 2005) in that the STS was preferentially activated by object-incongruent gaze shifts. Although it is difficult to speculate on the underlying cognitive mechanisms for such an object-incongruence effect, one possibility is that it reflects a

violation of expectation for human action. In other words, both adults and infants might have expected object-directed referential gaze shifts, and an object-incongruent gaze shift would have violated such expectation. This violation may have recruited additional cognitive processing. This account implies that these neural effects are sensitive to more general processing of human action. For example, Pelphrey, Morris and McCarthy (2004) reported that the STS is preferentially activated when participants watch another human perform non-object-directed reaching than when they perform object-directed reaching. Further studies will be required to test whether the adult N330 and/or the infant N290 are sensitive to the relations between reaching direction and the location of the object being reached for.

Only in infants did we observe anterior ERP components (anterior N200 and N400) that were higher amplitude in response to the perception of object-congruent gaze shifts. These frontal ERP components may reflect the recruitment of other parts of the social brain network. In an earlier study with adults, Sabbagh, Moulson and Harkness (2004) reported a frontal negativity (N270-N400), which was associated with the mental state judgments. In addition, ventro-medial prefrontal cortex, as well as STS and FFG, are elements of the cortical social brain network (Adolphs, 2003; Johnson et al., 2005), and may be involved in the attribution of the mental states (Frith & Frith, 1999; Saxe, Carey & Kanwisher, 2004). Interestingly, the anterior effect in the infant ERPs preceded the posterior component (N290), which seems to be functionally similar to adult N330. This temporal order raises the possibility that the anterior negativity is related to the hypothesised rapid sub-cortical route for face processing (Johnson, 2005). There are several possible reasons why only infants, and not adults, showed this frontal ERP effect.

The first possibility is that this effect reflects a developmental change in brain anatomy. In contrast to the present adult ERP results, Pelphrey et al. (2005) reported that in addition to the posterior networks including the STS, parts of frontal lobe including middle frontal gyrus (BA9), precentral gyrus (BA44) and cingulate gyrus (BA32) discriminated object-congruent and object-incongruent gaze. Other studies also reported that the prefrontal cortex is preferentially recruited by perceived direct gaze in adults (Kampe, Frith & Frith, 2003; Schilbach, Wohlschlaeger, Kraemer, Newen, Shah, Fink & Vogeley, 2006). Thus, it is possible that adults as well as infants recruit frontal cortex structures in the encoding of the referential information from the gaze, but that the scalp recorded ERP is insufficiently sensitive to detect this activity within the adult brain.

The second possibility is that the infant brain is less 'specialized' (Johnson et al., 2005) and that widespread cortical circuits are recruited in response to perceived gaze shifts. For example, the N290 ERP component in 4-month-old infants, but not the N170 in adults, shows differential amplitudes in response to faces with direct and averted eye gaze (Grice et al. 2005, Farroni et al. 2002). It is also possible that in the less specialized infant brain, the referential information of the gaze is encoded in broader cortical and subcortical circuits than in more specialized adult brain.

A third possibility is that the early prefrontal activity reflects the detection of communicative cues in infants, but not in adults, in the present experimental settings. Two points support this interpretation. Firstly, activity in prefrontal cortex is known to be modulated by the communicative cues including eye contact and hearing one's name called (Kampe et al., 2003). Secondly, this account may explain the greater amplitudes we observed for object-congruent gaze shifts in the anterior N200 and the N400. Specifically, object-congruent gaze signals communicative-referential context more than object-incongruent gaze, since the former refers to the object and is informative to the observer (Csibra & Gergely, 2006). Thus, it is possible that

greater amplitudes in the anterior N200 and the N400 reflect infants' detection of, and engagement in, the communicative context. This conclusion converges with our previous behavioural findings (Senju et al., 2006) that infants look longer to object-congruent gaze shifts compared to object-incongruent gaze shifts when they are preceded by eye contact. The lack of prefrontal activity in the adults may reflect the fact that direct gaze from a computer monitor may not be sufficient for adults to interpret the situation as true 'communication'. It is possible that a more socially engaging stimulus or task would also activate the anterior part of social brain network in adults (Kampe et al., 2003, Sabbagh et al., 2004, Schilbach et al., 2006).

The present study can be extended to the study of the atypical development of theory of mind in individuals with ASD (Baron-Cohen, 1995; Frith, 2001). Pelphrey et al. (2005) revealed that adults with ASD did not show differential STS activity for gaze-object congruency, which suggests that they did not encode referential information from others' gaze. The impairment in referential understanding of gaze shift is present in ASD from childhood (Baron-Cohen, Baldwin & Crowson, 1997). Since the present study found that the encoding of referential information from gaze appears as early as 9 months in typical development, it is possible that infants who later develop into ASD will show the lack of, or reduced, ERP signatures in response to the gaze-object congruency. Further work will be required to investigate the early development of referential understanding of gaze in infants at high risk for ASD.

References

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, **4**, 165-178.

Bach, M., & Ullrich, D. (1994). Motion adaptation governs the shape of motion-evoked cortical potentials. *Vision Research*, **34**, 1541-1547.

Bach, M., & Ullrich, D. (1997). Contrast dependency of motion-onset and pattern-reversal VEPs: interaction of stimulus type, recording site and response component. *Vision Research*, **37**, 1845-1849.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive*, **13**, 513-552.

Baron-Cohen, S. (1995). *Mindblindness : An essay on autism and theory of mind*. Cambridge, MA : MIT Press.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, **68**, 48-57.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "Language of the Eyes"? Evidence from normal adults, and adults with autism or Asperger Syndrome. *Visual Cognition*, **4**, 311-331.

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, **17**, 613-620.

Bloom, P. (2000). *How children learn the meaning s of words*. Cambridge, MA : MIT Press.

Csibra, G., & Gergely, G. (2006). Social learning and social cogniton: The case for pedagogy. In Y. Munakata & M. H. Johnson (Eds.), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI* (pp. 249-274). Oxford: Oxford University Press.

de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related

potentials during infancy: A review. International Journal of Psychophysiology, 51, 45-58.

de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, **14**, 199-209.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, **6**, 509-540.

Eimer, M. (1994). An ERP study on visual spatial priming with peripheral onsets. *Psychophysiology*, **31**, 154-163.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of National Academy of Science (USA)*, **99**, 9602-9605.

Farroni, T., Johnson, M. H., & Csibra, G. (2004a). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, **16**, 1320-1326.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004b). Gaze Following in Newborns. *Infancy*, **5**, 39-60.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Refrexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review*, **5**, 490-495.

Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, **32**, 969-979.

Frith, C. D., & Frith, U. (1999). Interacting minds-A biological basis. *Science*, **286**, 1692-1695.

Gibson, J. J., & Pick, A. D. (1963). Perception of another person's looking behavior. *American Journal of Psychology*, **76**, 386-394.

Grice, S. J., Halit, H., Farroni, T., Baron-Cohen, S., Bolton, P., & Johnson, M. H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*, **41**, 342-353.

Halit, H., Csibra, G., Volein, A., & Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*, **45**, 1228-1234.

Hirai, M., & Hiraki, K. (2005). An event-related potentials study of biological motion perception in human infants. *Cognitive Brain Research*, **22**, 301-304.

Hirai, M., Senju, A., Fukushima, H., & Hiraki, K. (2005). Active processing of biological motion perception: an ERP study. *Cognitive Brain Research*, **23**, 387-396.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, **3**, 80-84.

Itier, R. J., & Taylor, M. J. (2004). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *Neuroimage*, **21**, 1518-1532.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, **6**, 766-774.

Johnson, M. H., de Haan, M., Oliver, A., Smith, W., Hatzakis, H., Tucker, L. A., & Csibra, G. (2001). Recording and analyzing high density ERPs with infants using the Geodesic Sensor Net. *Developmental Neuropsychology*, **19**, 295-323.

Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., de Haan, M., Tucker, L. A., Baron-Cohen, S., & Richards, J. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Development and Psychopathology*, **17**, 599-619.

Jokisch, D., Daum, I., Suchan, B., & Troje, N. F. (2005). Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. *Behavioral Brain Research*, **157**, 195-204.

Kampe, K. K., Frith, C. D., & Frith, U. (2003). "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *Journal of Neuroscience*, **23**, 5258-5263.

Kleinke, C. (1986). Gaze and eye-contact: A research review. *Psychological Bulletin*, **100**, 78-

100.

Klucharev, V., & Sams, M. (2004). Interaction of gaze direction and facial expressions processing: ERP study. *Neuroreport*, **15**, 621-625.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature*, **387**, 767-768.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of Human Evolution*, **40**, 419-435.

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, **6**, 541-567.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters*, **253**, 147-150.

Meeren , H. K., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of National Academy of Science (USA)*, **102**, 16518-16523.

Mosconi, M. W., Mack, P. B., McCarthy, G., & Pelphrey, K. A. (2005). Taking an 'intentional stance' on eye-gaze shifts: A functional neuroimaging study of social perception in children. *NeuroImage*, **27**, 247-252.

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the intentions of others: The perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *Journal of Cognitive Neuroscience*, **16**, 1706-1716.

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain*, **128**, 1038-1048.

Pelphrey, K. A., Singerman, J. D., Allison, T., & McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. *Neuropsychologia*, **41**, 156-170.

Pelphrey, K. A., Viola, R. J., & McCarthy, G. (2004). When strangers pass: processing of mutual and averted social gaze in the superior temporal sulcus. Psychological Science, 15, 598-603.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, **18**, 2188-2199.

Puce, A., Smith, A., & Allison, T. (2000). ERPs evoked by viewing facial movements. *Cognitive Neuropsychology*, **17**, 221-239.

Reid, V. M., Hoehl, S., & Striano, T. (2006). The perception of biological motion by infants: An event-related potential study. *Neuroscience Letters*, **395**, 211-214.

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. Neuroreport, 15, 2553-2555.

Ricciardelli, P., Ro, T., & Driver, J. (2002). A left visual field advantage in perception of gaze direction. *Neuropsychologia*, **40**, 769-777.

Sabbagh, M. A., Moulson, M. C., & Harkness, K. L. (2004). Neural correlates of mental state decoding in human adults: an event-related potential study. *Journal of Cognitive Neuroscience*, **16**, 415-426.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual Reviews of Psychology*, **55**, 87-124.

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., & Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions.

Neuropsychologia, **42**, 1435-1446.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia,* **44**, 718-730.

Schuller, A., & Rossion, B. (2001). Spatial attention triggered by eye gaze increases and speeds up early visual activity. *Neuroreport*, **12**, 2381-2386.

Senju, A., Csibra, G., & Johnson, M. H. (2006). Understanding the referential nature of looking: Infants' preference for object-directed gaze. Manuscript submitted for publication.

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*, **12**, 1474-1496.

Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry*, **45**, 445-458.

Senju, A., Yaguchi, K., Tojo, Y., & Hasegawa, T. (2003). Eye contact does not facilitate detection in children with autism. *Cognition*, **89**, B43-B51.

Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y., & Mazoyer, B. (2002). Neural correlates of woman face processing by 2-month-old infants. *Neuroimage*, **15**, 454-461.

Webb, S. J., Long, J. D., & Nelson, C. A. (2005). A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science*, **8**, 605-616.

Wicker, B., Michel, F., Henaff, M. A., & Decety, J. (1998). Brain regions involved in the perception of gaze: a PET study. *NeuroImage*, **8**, 221-227.