Audiovisual temporal fusion in 6-month-old infants☆

Franziska Kopp*

Max Planck Institute for Human Development, Berlin, Germany

ARTICLE INFO

ABSTRACT

The aim of this study was to investigate neural dynamics of audiovisual temporal fusion processes in 6-month-old infants using event-related brain potentials (ERPs). In a habituation-test paradigm, infants did not show any behavioral signs of discrimination of an audiovisual asynchrony of 200 ms, indicating perceptual fusion. In a subsequent EEG experiment, audiovisual synchronous stimuli and stimuli with a visual delay of 200 ms were presented in random order. In contrast to the behavioral data, brain activity differed significantly between the two conditions. Critically, N1 and P2 latency delays were not observed between synchronous and fused items, contrary to previously observed N1 and P2 latency delays between synchrony and perceived asynchrony. Hence, temporal interaction processes in the infant brain between the two sensory modalities varied as a function of perceptual fusion versus asynchrony perception. The visual recognition components Pb and Nc were modulated prior to sound onset, emphasizing the importance of anticipatory visual events for the prediction of auditory signals. Results suggest mechanisms by which young infants predictively adjust their ongoing neural activity to the temporal synchrony relations to be expected between vision and audition.

© 2014 The Author. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Temporal synchrony is one of the strongest binding cues in multisensory perception (King, 2005; Spence and Squire, 2003). Yet, in order to perceive simultaneity in a bimodal stimulus, perfect temporal synchrony of two sensory streams is not required. For audition and vision, the concept of a temporal window of integration was proposed in which auditory and visual input are pulled into temporal alignment to result in a fused, simultaneous percept (e.g., Fendrich and Corballis, 2001; Lewkowicz, 1996, 2000; Van Wassenhove et al., 2007; Vatakis et al., 2007). This temporal window was identified as being flexible and depending on a variety of parameters such as complexity of stimuli, familiarity and experience, or repeated asynchrony presentation (e.g., Dixon and Spitz, 1980; Fujisaki et al., 2004; Navarra et al., 2005, 2010; Petrini et al., 2009; Powers et al., 2009; Vatakis and Spence, 2006). The temporal range has a certain degree of variability across individuals (Stevenson et al., 2012) and appears to undergo changes across the lifespan (Hillock et al., 2011; Lewkowicz, 1996, 2010).

From early on, human infants are sensitive to intersensory temporal synchrony relations (e.g., Bahrick, 1983; Dodd, 1979; Hollich et al., 2005; Lewkowicz, 1986, 1992; Lewkowicz et al., 2008, 2010; Spelke, 1979). Multisensory capacities improve and responsiveness to complex multisensory temporal relations increases in the first months of life (e.g., Bahrick, 1987; Lewkowicz, 2000; Lewkowicz

et al., 2008). For infants, Lewkowicz demonstrated that the window of perceptual fusion is larger than in adults, both for simple, abstract and for speech stimuli (cf. Dixon and Spitz, 1980; Lewkowicz, 1996, 2010; Vatakis and Spence, 2006). Hence, one may assume that temporal fusion processes undergo developmental changes from infancy to adulthood. This study is one of the first to document neural dynamics of audiovisual temporal fusion in infants, investigating an audiovisual temporal disparity within the temporal window of integration.

In adults, brain activity modulations related to audiovisual temporal synchrony perception were found in large-scale neural networks (e.g., Bushara et al., 2001; Dhamala et al., 2007; Macaluso et al., 2004; Stevenson et al., 2010). Evaluation of synchronous versus asynchronous presentation was shown to involve different networks (superior colliculus, anterior insula, anterior intraparietal sulcus) than successful perceptual fusion does (Heschl's gyrus, superior temporal sulcus, middle intraparietal sulcus, inferior frontal gyrus; Miller and D'Esposito, 2005; see also Stevenson et al., 2011), suggesting a functional dissociation between the mechanisms of physical synchrony and subjective simultaneity perception. Studies using event-related brain potentials (ERPs) have demonstrated that the auditory components N1 and P2 were sensitive to bimodal audiovisual versus unimodal auditory stimulation (e.g., Besle et al., 2009). Importantly, these modulations depended on the salience of the visual input and on anticipatory visual motion (Stekelenburg and Vroomen, 2007; Vroomen and Stekelenburg, 2009; Van Wassenhove et al., 2005) and were absent when the two sensory signals were not in synchrony (Pilling, 2009). Although the influence of temporal synchrony relations has been addressed in several studies (e.g., Talsma et al., 2009; Vroomen and Stekelenburg, 2009), electrophysiological dynamics of audiovisual perceptual fusion in adults merit further investigation.

Research on neural processes related to multisensory perception in infants is still developing. In a series of experiments on audiovisual perception, Hyde and colleagues found that, in contrast to adult ERP data, the auditory component P2 was not sensitive to multisensory versus unisensory presentation of circles and tones in 3-month-olds, but was modulated by manipulations of dynamic versus static faces and of audiovisual congruency in speech stimuli in 5-month-olds (Hyde et al., 2010, 2011). Visual recognition dynamics are reflected in the infant components Nc and Pb. Nc is a negative peak between 400 and 700 ms after stimulus onset, which has been related to mechanisms of attention and memory (e.g., Ackles and Cook, 2007; Kopp and Lindenberger, 2011, 2012; Reynolds and Richards, 2005). Pb is a smaller positive deflection peaking between 250 and 450 ms. It has been observed to be modulated by expectancy processes and the relevance of stimuli (e.g., Karrer and Monti, 1995; Kopp and Dietrich, 2013; Kopp and Lindenberger, 2011, 2012; Nikkel and Karrer, 1994).

A recent study at our lab (Kopp and Dietrich, 2013) investigated audiovisual synchrony and asynchrony perception in 6-month-old infants using ERP. Movies of a person clapping her hands were presented with visual and auditory input in synchrony in one condition and a visual delay of 400 ms in the other condition. Infants discriminated the 400-ms asynchrony behaviorally in a habituation-test task. ERPs revealed latency shifts of the auditory N1 and P2 between asynchronous and synchronous events, although the auditory input occurred at the same point in time in both experimental conditions. The magnitude of this shift indicated a temporal interaction between the two modalities. It was hypothesized that these latency delays in the infant auditory ERP components might be indicators for the emergence of an asynchronous percept on the behavioral level. Importantly, neural processing was already affected prior to the auditory onset, suggesting anticipatory mechanisms as to the timing of the two sensory modalities. Nc latency shifts implied an attentional shift in time between synchrony and asynchrony. Moreover, the polarity of Pb was reversed, being related to predictive processes as to audiovisual temporal synchrony relations prior to sound onset (for details see Kopp and Dietrich, 2013).

To date there is little insight into the emergence of multisensory percepts in infants and underlying neural activity. While behavioral performance indicates simultaneity perception both in physically synchronous and perceptually fused stimuli, differential neural processing is very likely (Miller and D'Esposito, 2005; Stevenson et al., 2011). The aim of the present study was to investigate neural dynamics within the temporal window of integration, that is, when the visual delay is smaller than the asynchrony tolerance. The paradigm used in Kopp and Dietrich (2013) was adapted for this purpose. First, a standardized infant-controlled habituation-test paradigm was applied. Children were tested for discrimination of a visual delay of 200 ms in audiovisual stimuli, an asynchrony known to correlate with simultaneity perception in infants (as proven by extensive piloting in our lab and by findings of Lewkowicz, 1996). Then, EEG activity was assessed in response to audiovisually synchronous stimuli and to stimuli in which the visual stream was delayed by 200 ms with respect to the auditory stream.[1]

It was predicted that infants would not be able to detect the 200-ms asynchrony behaviorally. However, following adult neuroimaging studies (e.g., Miller and D'Esposito, 2005; Stevenson et al., 2011), neural activity should differ between audiovisually synchronous and perceptually fused items. According to the discussion by Kopp and Dietrich (2013) that significant latency delays in the infant auditory components N1 and P2 might be an indicator for the emergence of an asynchronous percept on the behavioral level, one would predict these latency differences to disappear in the present paradigm, as the 200-ms

---

[1] As in the Kopp and Dietrich (2013) study, only the content of the visual input was delayed, while keeping both the video and audio onset times and durations identical between the two experimental conditions. This setup avoided differences due to attentional shifts as orienting responses to stimulus onsets and offsets during the presentation, and attentional competition between the two sensory modalities (e.g., Talsma et al., 2010).

asynchrony would not elicit a behaviorally asynchronous percept in the infant sample.

2. Materials and methods

2.1. Participants

The participants were 6-month-old infants for comparability with the Kopp and Dietrich (2013) study. Sixty-nine infants and their parents were invited to the Baby Laboratory at the Max Planck Institute for Human Development, Berlin. All infants were born full-term ($\geq$38th week), with birth weights of 2500 g or more. According to the parents' evaluation, all participants were free of neurological diseases, and had normal hearing and vision. The institute's Ethics Committee approved of this study. Informed written consent was obtained from the infants' caregivers.

Five children were excluded from the initial sample due to an experimental error ($n = 2$) or due to the fatigue criterion (see below; $n = 3$) in the habituation task. For EEG analysis, another 28 participants were excluded due to failure to reach the minimum requirements for adequate ERP averaging, because of excessive fussiness, movement artifacts, or insufficient visual fixation. The final sample included 36 infants (22 girls, 14 boys) with an age range between 170 and 190 days ($M = 178.7$ days, $SD = 6.3$ days).

2.2. Habituation task

2.2.1. Stimuli

Videos like the stimuli used in a previous study were presented (see Materials and Methods in Kopp and Dietrich, 2013). These videos consisted of sequences of a female person clapping her hands at a rate of approximately 1 per second (mean deviation of clapping intervals: ± 21.7 ms). The sampling rate of the recordings was 25 frames per second. The video showed the female's face, her shoulders, and her hands in front of the face (see Fig. 1).

Two stimuli were created, with durations of 30 s each. The first was an unaltered synchronous hand-clapping sequence (AVsynch). The other was a physically asynchronous stimulus (AVdelay; Fig. 1)[2] in which the visual stream was delayed by 200 ms (5 frames) with respect to the auditory stream using the editing software Adobe Premiere® 6.0. Loudness was 53 dB SPL (A) in both stimuli.

2.2.2. Procedure

Infants sat on their parent's lap in a quiet, acoustically shielded room, in an area surrounded by white curtains to prevent visual distraction. A computer screen was placed at a distance of 80 cm in front of the infant, as well as two loudspeakers to the left and to the right side of the screen with a distance of approximately 100 cm to the infant's head. One camera was placed above and behind the screen to capture the infant's gazing behavior during the experiment. The video recordings and the stimulus presentations were connected to a mixer leading to a time-synchronized split-screen image of infant and stimuli.

The procedure of the infant-controlled habituation paradigm included a pre-test, habituation phase, test phase, and a post-test. The pre- and post-test served as controls for alertness before habituation trials and after test trials. This behavioral control was implemented to increase the validity of the habituation data because a very low level of initial and/or terminal attention could be an indicator for a state of generalized fatigue, which in turn could affect responsiveness during the task. For this purpose, sequences of an age-appropriate child movie were presented for as long as the child looked at them or for a maximum duration of 30 s. Gaze durations of less than 5 s in the pre- and/or post-test led to the exclusion of the infant (fatigue criterion).

In the habituation phase, the synchronous stimulus was presented repeatedly. Each of the habituation trials lasted as long as the child looked at the stimulus or for a maximum duration of 30 s. An experimenter monitored the infant's gaze. Whenever the child looked away from the screen, another sequence of the child movie was presented to attract attention back to the screen. Gaze away was defined as lasting at least 1 s and including a head movement. When the infant continued to look at the monitor for 5 s or longer, the experimenter switched to the next habituation trial. Infants' gaze duration was recorded. The habituation criterion was reached when the mean gaze duration to the last three habituation trials was smaller than 50% of the mean gaze duration to the first three habituation trials.

In the test phase, the familiar synchronous stimulus (AVsynch) and the novel test stimulus (AVdelay) were presented to test for recovery of interest. Again, the duration of the trials was as long as the child looked at the screen or for a maximum of 30 s. Sequences of the child movie appeared to attract attention to the monitor, and when the infant looked at them for 5 s or longer, the experimenter switched to the next test trial. The order of test trials did not vary across the children. All infants saw and heard AVdelay after the last habituation trial, and then AVsynch before the post-test trial was presented.

2.3. EEG task

2.3.1. Stimuli

Two stimuli were created by cutting a sequence in which the inter-clap interval was precisely 1000 ms. The synchronous stimulus consisted of four hand-clapping movements in synchrony with the corresponding hand-clapping sound (duration: 4000 ms). Auditory onsets were at 474 ms, 1474 ms, 2474 ms, and 3474 ms after video onset. The stimulus occupied a visual angle of 8.7° when the female had extended both arms, and 7.1° when her palms had clapped. The hand-clapping sounds had equalized sound pressure level (48 dB SPL (A)), with a rise time of about 10 ms and a fade-out period of about 170 ms.

In the AVdelay stimulus, the content of the visual stream was delayed by 200 ms with respect to the auditory stream. That is, the video showed the hands of the female clapping

---

[2] In this article, the label *AVdelay* is preferred as the term *asynchronous* may imply behavioral asynchrony perception which is actually absent in fused stimuli.
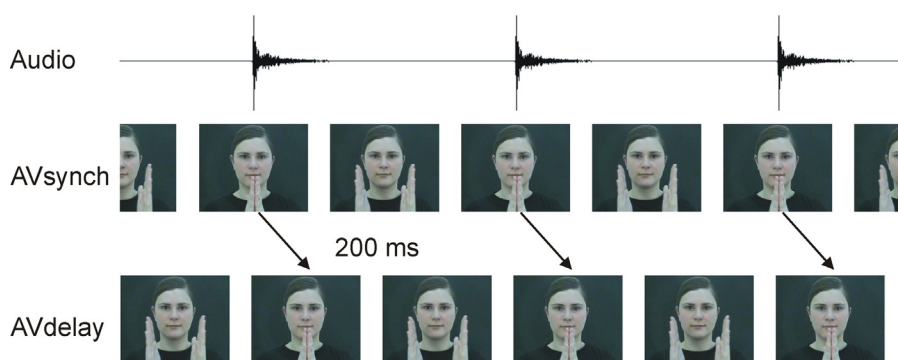
Fig. 1. Stimulus material of the habituation and the EEG paradigm. Videos were created showing a female clapping her hands at a rate of 1 per second. In AVsynch stimuli, the visual and the auditory signals were temporally aligned. In AVdelay stimuli, the video stream was delayed by 200 ms with respect to the auditory stream.

at 674 ms, 1674 ms, 2674 ms, and 3674 ms after stimulus onset. The video sequence of the continuous movement of 200 ms was added at the beginning and the video sequence of 200 ms at the end was cut out. In contrast, the time course of the auditory stream did not differ from the synchronous stimulus, with sound onsets at 474 ms, 1474 ms, 2474 ms, and 3474 ms.

2.3.2. Procedure

The surroundings, experimental setting, and video recording were the same as in the habituation experiment. EEG assessment took place right after the habituation task. To minimize EEG artifacts, parents were briefed not to talk to the child and to avoid any movements.

The experimental design included repeated presentations of AVsynch and AVdelay stimuli in random order. In both conditions, each trial started with an alternating sequence of an animated child movie (as used during the habituation paradigm) to direct and maintain the child's attention on the screen. For each trial, these animated movie sequences were randomly selected out of 20 sequences with varying durations between 3000 and 6000 ms. Next, in both experimental conditions, a static photo of the female was presented for 1000 ms as a baseline. For this purpose, a snapshot of the subsequent clapping stimulus was adapted. The hands of the woman were removed and replaced by the background of the picture. Following the baseline, either the AVsynch or the AVdelay stimulus was presented for 4000 ms.

Thus, trials lasted between 8000 and 11,000 ms. No more than three AVsynch or three AVdelay stimuli were presented consecutively. Whenever the infant became fussy or did not look at the screen any longer, sequences of the animated child movie were presented again, and then stimulus presentation continued. The average frequency of presentation of these sequences during a session was 5.2. This procedure was independent of the stimulus type presented, and the number of unattended trials before the presentation was variable. The session ended when the infant's attention could no longer be attracted to the screen. Within the session, a maximum of 90 trials of AVsynch and 90 trials of AVdelay stimuli were presented. Infants saw and heard on average 43.7 AVsynch

and 36.9 AVdelay stimuli (AVsynch vs. AVdelay: $t(35) = 1.5$, $p = .140$).

2.3.3. EEG acquisition and analysis

EEG was continuously recorded at 32 active electrodes. Signals were recorded with a sampling rate of 1000 Hz and amplified by a Brain Vision amplifier. The reference electrode was placed at the right mastoid, and the ground electrode at AFz. Signals at FP1 and FP2 were monitored to check for vertical eye movements, and signals at F9 and F10, for horizontal eye movements. Impedances were kept below 20 kΩ.

As in the habituation experiment, a video camera captured the child's gazing behavior. Video recordings, stimulus presentations, and event markers of each trial sequence were connected to a mixer leading to a time-synchronized split-screen image. All trials in which the infant did not look at the screen were excluded from further EEG analysis.

Defect or noisy channels were interpolated using a spherical spline algorithm (Perrin et al., 1990). EEG was re-referenced off-line to linked mastoids. A bandpass filter was set off-line between 0.5 and 20 Hz. Previous results (Bristow et al., 2009; Hyde et al., 2010; Kopp and Dietrich, 2013) indicate that effects due to audiovisual temporal synchrony perception occur early after stimulus onset. Therefore the final EEG analysis segments were set to the first 1000 ms after video onset with the sound occurring at 474 ms. ERP epochs comprised a 200 ms baseline before video onset followed by 1000 ms of video presentation. Artifacts due to eye or body movements or external sources were automatically discarded when voltage exceeded ±120 μV. In addition, EEG signals were inspected visually to scan for and reject artifacts. A baseline correction to the 200 ms pre-stimulus baseline was performed. Finally, individual averages (ERP) and grand averages across subjects were calculated. For ERP analysis, infants contributed on average 23.9 trials, with AVsynch stimuli and 19.7 trials with AVdelay stimuli to their grand averages (AVsynch vs. AVdelay: $t(35) = 1.5$, $p = .131$). Peaks of the ERP components of interest (see Section 3.2.1 below) were identified using a semiautomatic procedure: First, local maxima were automatically detected for the predefined intervals and polarity, separately for each channel,
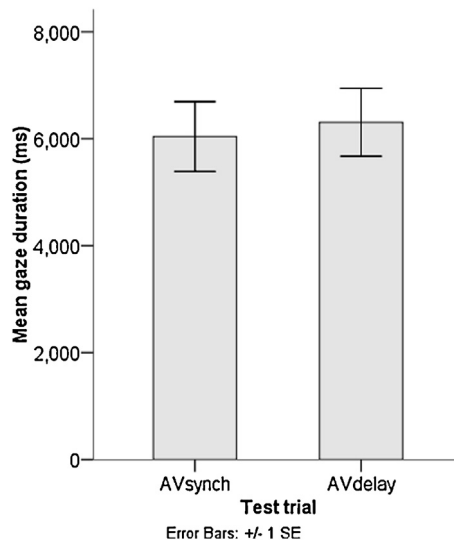
Fig. 2. Results of the behavioral habituation paradigm. Similar gaze durations to the familiar synchronous stimulus (AVsynch) and to the novel AVdelay stimulus suggest a lack of recovery of interest after habituation to synchronous stimuli.

taking into account topographic variability. Second, this procedure included a fine adjustment of the peaks via visual inspection. Finally, peak amplitudes and latencies were exported for statistical analysis (see Section 3.2).

3. Results

3.1. Behavioral data

Mean gaze duration of the test trials is depicted in Fig. 2. Infants reached habituation to the synchronous stimulus on average after 8.6 presentations ($SD = 3.7$). High gaze duration to the animated child movie in the pre-test ($M = 23912.4$ ms, $SD = 8096.1$ ms) and post-test ($M = 21740.6$ ms, $SD = 8611.5$ ms) indicated high alertness during the habituation-test procedure.

A paired-samples t test revealed that the looking time to AVdelay ($M = 6307.4$ ms, $SD = 3821.7$ ms) did not differ significantly from the looking time to AVsynch ($M = 6040.7$ ms, $SD = 3923.8$ ms), $t(35) = 0.40$, $p = .688$. This result suggests that infants did not show recovery of interest to AVdelay after habituation to the synchronous stimulus. It is important to note that the link between the lack of a significant difference in gaze duration between familiar and novel test trials and the lack of asynchrony discrimination is not mandatory. Other factors, such as attention and memory processes, prior knowledge, and experience, may also influence infant gazing behavior in a habituation paradigm (for a detailed discussion, see for example Aslin, 2007, or Colombo and Mitchell, 2009). However, the present stimulus material was piloted extensively in the lab and significant recovery of interest was observed for audiovisual asynchronies of 280 ms or more in group level data (see also discrimination of 400-ms asynchrony in 6-month-olds in Kopp and Dietrich, 2013). Furthermore, the data are in line with previous research showing

audiovisual asynchrony discrimination in non-speech stimuli in 2- to 8-month-old infants only in asynchronies of 300 ms or more (Lewkowicz, 1996). In other words, it is fair to assume that the physical asynchrony of 200 ms corresponds to temporally fused, simultaneity perception in the 6-month-old infants of the present sample.

3.2. EEG data

3.2.1. Overview

As in the previous study (Kopp and Dietrich, 2013), the ERP components N1 and P2 were elicited as responses to the auditory onset (at 474 ms; see Fig. 3). N1 peaked on average 149.8 ms and P2 on average 294.4 ms after sound onset (i.e., 623.8 ms and 768.4 ms, respectively, after video onset). Although N1 activity partly overlapped with prior responses to the visual stimulation, the peak was reliably identified (see Section 2.3.3) and was therefore analyzed as the most negative deflection in the interval between 550 and 700 ms after video onset. The analysis interval for P2 was between 700 and 900 ms after video onset.

Earlier in the EEG epoch, the infant ERP components Pb (interval: 300–425 ms) and Nc (interval: 425–550 ms) were observed as being differentially activated in AVsynch and AVdelay visual stimulation (Fig. 3). For statistical analysis, repeated-measures analyses of variance (ANOVA) were performed. The significance level of $\alpha = .05$ was Bonferroni-adjusted to control for multiple comparisons in each ANOVA. Separate ANOVAs were performed for the different dependent variables, these were peak amplitude and peak latency of N1, P2, Pb, and Nc. Means and standard deviations are reported in Table 1.

Visual data inspection suggests that all ERP components of interest were elicited across the entire scalp, displaying some topographical variability, with differences between the two experimental stimuli primarily in anterior electrodes (Figs. 3 and 4). This topographical pattern resembles the previous data by Kopp and Dietrich (2013), suggesting a similar anterior–posterior differentiation as identified in that study. Therefore statistical analyses including a test of topographic variability and of stimulus effects across all electrodes were performed here first. Next, specific effects of AVsynch versus AVdelay stimuli were explored further in anterior electrodes taking into account further possible lateralization effects.

3.2.2. Topography and stimulus effects

ANOVAs included the within-subjects factors Stimulus (AVsynch vs. AVdelay) and Region with the following regions defined by electrode lines from anterior to posterior: frontal (F7, F3, Fz, F4, F8), fronto-central (FC5, FC1, FC2, FC6), central/temporal (T7, C3, Cz, C4, T8), centro-parietal (CP5, CP1, CP2, CP6), parietal (P7, P3, Pz, P4, P8), and parieto-occipital (PO9, O1, O2, PO10).

3.2.2.1. Auditory-evoked components.
Analyses revealed a main effect of Region on P2 amplitude, $F(5, 175) = 5.47$, $p < .001$, $\eta^2 = .14$. More specifically, P2 peak amplitudes were on average more positive in anterior than in posterior electrodes (significant post hoc pairwise comparisons: frontal vs. parietal, parieto-occipital; fronto-central vs.
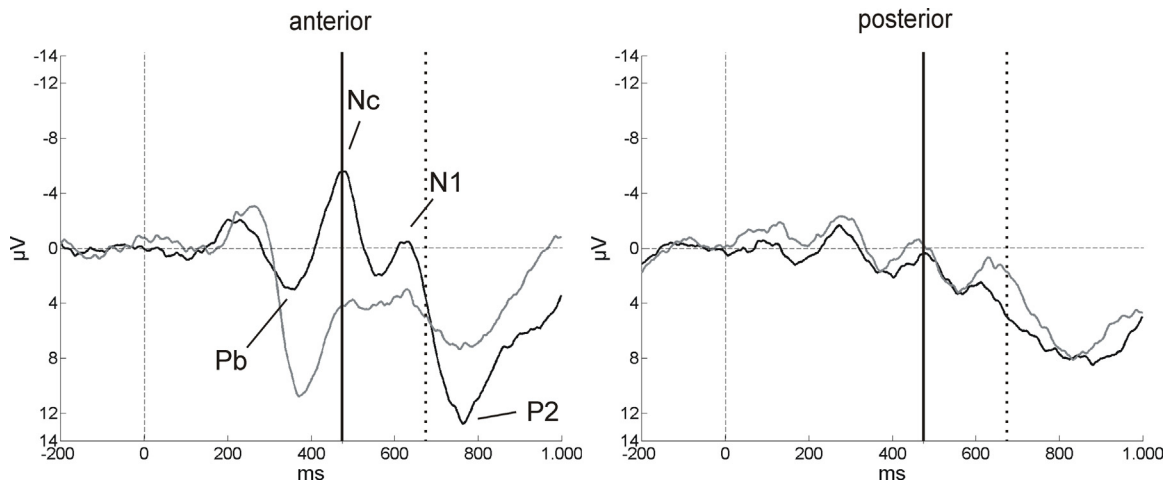
Fig. 3. ERP components of interest in the first 1000 ms after video onset. The ERP signals elicited by AVsynch (black line) and AVdelay stimuli (gray line) were averaged across anterior and across posterior electrodes for the purpose of illustration. Sound onset at 474 ms is indicated by vertical, solid lines. The visual clapping of the hands occurred at 674 ms in AVdelay as indicated by vertical, dotted lines.

parietal, parieto-occipital; for a detailed report of post hoc pairwise statistics see supplementary material). Additionally, a main effect of Region was found on N1 latency, $F(5, 175) = 4.59$, $p = .001$, $\eta^2 = .12$, and on P2 latency, $F(5, 175) = 2.72$, $p = .022$, $\eta^2 = .07$, both with longer latencies at posterior sites (significant post hoc pairwise comparisons N1 latency: frontal vs. parietal, parieto-occipital; fronto-central vs. parietal, parieto-occipital; central/temporal vs. parietal, parieto-occipital; P2 latency: frontal vs. parieto-occipital; fronto-central vs. parieto-occipital; central/temporal vs. parieto-occipital).

Across all electrodes, differences due to the experimental manipulation (Stimulus) were not reliable in the auditory components. P2 amplitude was marginally smaller in AVdelay compared to AVsynch trials, $F(1, 35) = 3.38$, $p = .074$, $\eta^2 = .09$, but this difference did not reach statistical significance at the 5% level.

For the N1 component, a significant Region × Stimulus interaction was revealed. Negative responses were diminished in AVdelay as compared to AVsynch stimuli specifically in anterior electrodes, $F(5, 175) = 3.17$, $p = .009$, $\eta^2 = .08$.

3.2.2.2. Visual-evoked components. Significant main effects of Region were found on Pb amplitude, $F(5, 175) = 10.63$, $p < .001$, $\eta^2 = .23$, and on Nc amplitude, $F(5, 175) = 3.05$, $p = .012$, $\eta^2 = .08$. Pb amplitudes were on average more positive in anterior than in posterior electrodes, whereas Nc amplitude was less negative at posterior sites (significant post hoc pairwise comparisons Pb amplitude: frontal vs. centro-parietal, parietal, parieto-occipital; fronto-central vs. centro-parietal, parietal, parieto-occipital; central/temporal vs. centro-parietal, parietal, parieto-occipital; Nc amplitude: frontal vs. centro-parietal; fronto-central vs. parietal; central/temporal vs. parietal, parieto-occipital).

Table 1
Means and standard deviations of the analyzed ERP measures for AVsynch and AVdelay stimuli.

			All electrodes		Anterior electrodes	
			Mean	SD	Mean	SD
N1	Amplitude (μV)	AVsynch	−2.8	1.5	−4.2	1.9
		AVdelay	−0.9	1.5	−0.3	1.9
	Latency (ms)	AVsynch	620.7	6.0	619.1	6.1
		AVdelay	627.0	4.9	625.0	5.2
P2	Amplitude (μV)	AVsynch	14.0	1.7	15.9	2.1
		AVdelay	11.2	1.3	11.8	1.6
	Latency (ms)	AVsynch	762.2	8.1	760.6	8.0
		AVdelay	774.6	8.3	772.5	8.4
Pb	Amplitude (μV)	AVsynch	4.8	0.9	5.8	1.2
		AVdelay	8.9	1.2	13.8	1.7
	Latency (ms)	AVsynch	357.8	4.9	357.7	5.0
		AVdelay	376.3	4.7	377.6	4.8
Nc	Amplitude (μV)	AVsynch	−5.8	1.2	−7.8	1.5
		AVdelay	−2.6	1.5	−0.7	1.9
	Latency (ms)	AVsynch	474.8	4.1	474.8	3.7
		AVdelay	482.6	6.2	484.7	6.5

Note. Descriptives are collapsed across the factor Region in the analysis of all electrodes and across the factor Hemisphere in the analysis of anterior electrodes.
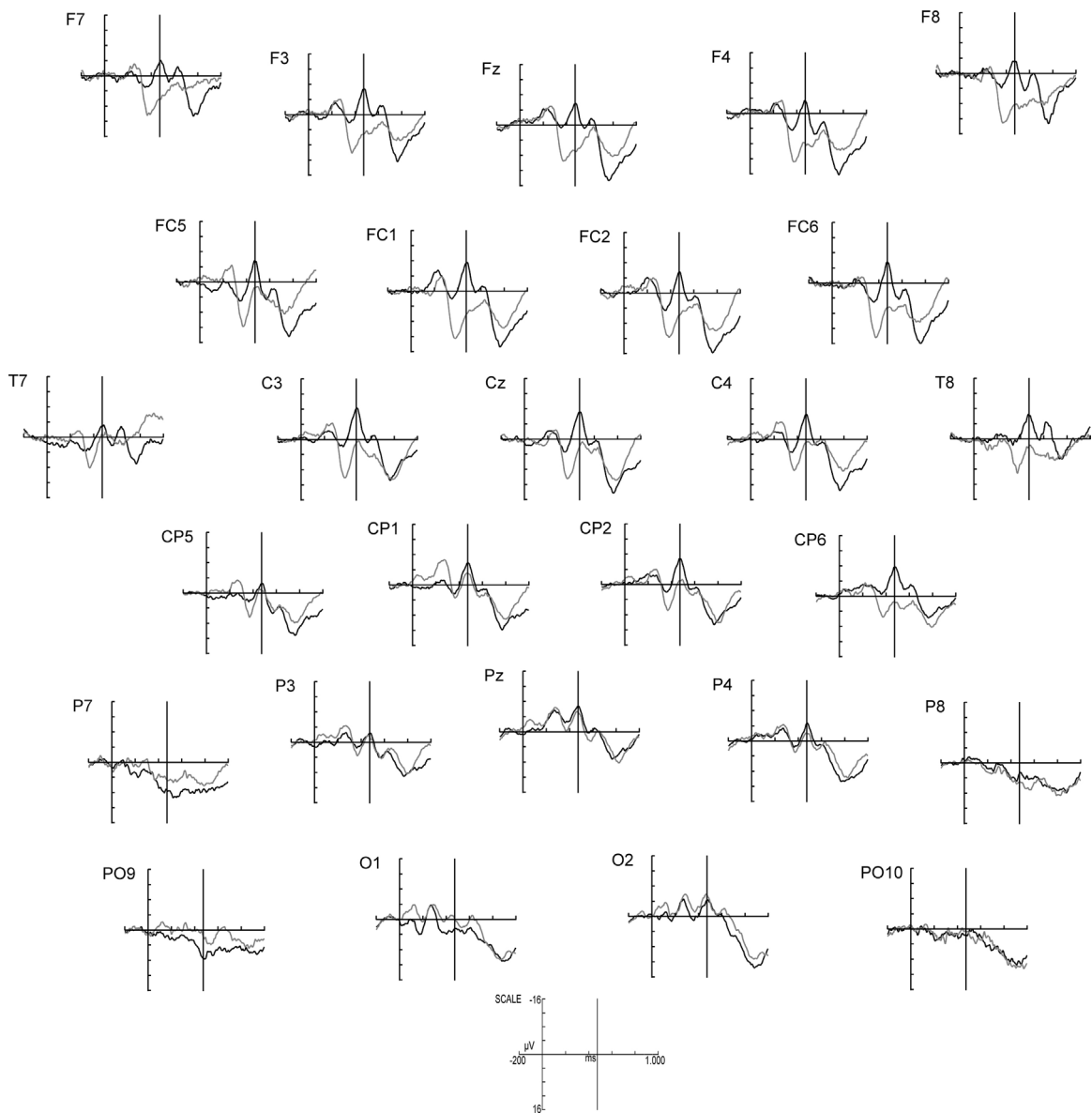
Fig. 4. ERPs elicited by AVsynch (black line) and AVdelay (gray line) stimuli in the first 1000 ms of stimulus presentation. Sound onset at 474 ms in both experimental conditions is indicated by a vertical, solid line at each electrode.

Stimulus had a main effect on Pb amplitude, $F(1, 35) = 8.91$, $p = .005$, $\eta^2 = .20$, with more positive voltages in AVdelay, and for Nc amplitude, $F(1, 35) = 5.02$, $p = .032$, $\eta^2 = .12$, with reduced negativity in AVdelay (Fig. 5). In addition, a main effect of Stimulus on Pb latency, $F(1, 35) = 6.85$, $p = .013$, $\eta^2 = .16$, indicated overall longer Pb latencies in AVdelay than in synchronous stimuli.

The main effects of Region and Stimulus on Pb and Nc amplitude were qualified by reliable Region × Stimulus interactions. Pb amplitude was larger in AVdelay than in AVsynch stimuli and Nc amplitude was diminished in AVdelay in anterior but not in posterior electrodes (Region × Stimulus on Pb amplitude: $F(5, 175) = 11.80$,

$p < .001$, $\eta^2 = .25$, and on Nc amplitude: $F(5, 175) = 10.78$, $p < .001$, $\eta^2 = .24$; see Fig. 5).

In sum, the Region main effects document an anterior–posterior differentiation in all analyzed ERP components (see Figs. 3 and 4), confirming the definition of anterior electrodes as a region of interest for more detailed analyses.

3.2.3. Anterior differentiation

Specific stimulus effects were explored in greater detail at frontal, fronto-central, and central/temporal sites (see also Kopp and Dietrich, 2013). ANOVAs were performed including the factors Stimulus (AVsynch vs. AVdelay) and Hemisphere with the following regions of interest:
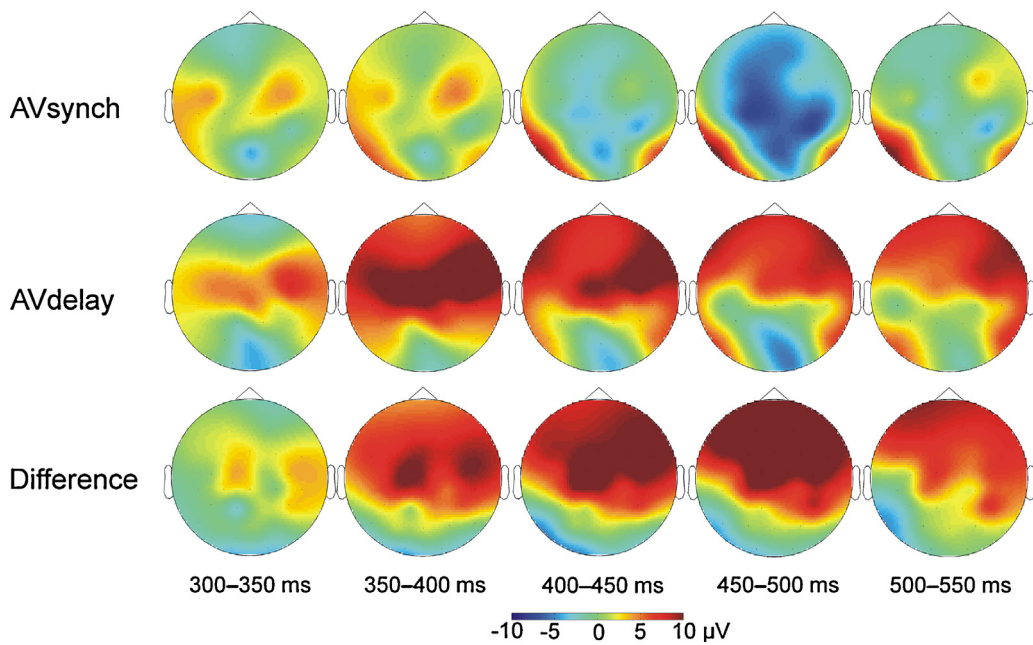
Fig. 5. Topographical maps for AVsynch and AVdelay stimuli and the difference (AVdelay–AVsynch) in the time window of Pb and Nc activity. Early significant differences between the two experimental conditions indicate perceptual fusion dynamics before sound onset, primarily at anterior sites.

left-hemisphere (F7, F3, FC5, T7, C3), midline (Fz, FC1, FC2, Cz), and right-hemisphere (F4, F8, FC6, C4, T8) electrodes.

3.2.3.1. Auditory-evoked components. An influence of the factor Hemisphere was evident only in the P2 responses. At midline electrodes, P2 voltage was higher overall and P2 peak latency was longer as compared to those at the left- and right-hemisphere locations (main effect Hemisphere on P2 amplitude: $F(2, 70) = 6.23$, $p = .003$, $\eta^2 = .15$, on P2 latency: $F(2, 70) = 3.32$, $p = .042$, $\eta^2 = .09$), however, these Hemisphere effects did not interact with Stimulus.

Differences between the two stimuli were observed in the auditory components in the anterior region of interest. Analyses showed a trend for diminished N1 amplitudes in AVdelay as compared to synchronous stimuli, $F(1, 35) = 3.98$, $p = .054$, $\eta^2 = .10$, without reaching significance at the 5% level. However, the main effect of Stimulus on P2 amplitude became significant at anterior electrodes, $F(1, 35) = 4.48$, $p = .041$, $\eta^2 = .11$, in contrast to the marginal effect across all electrodes at the scalp (see above Section 3.2.2.1). Accordingly, smaller P2 responses were found in AVdelay stimuli.

3.2.3.2. Visual-evoked components. Main effects of Stimulus on Pb amplitude, $F(1, 35) = 25.41$, $p < .001$, $\eta^2 = .42$, Pb latency, $F(1, 35) = 7.93$, $p = .008$, $\eta^2 = .18$, and Nc amplitude, $F(1, 35) = 15.22$, $p < .001$, $\eta^2 = .30$, were confirmed at higher significance levels at anterior sites. That is, compared to AVsynch stimuli, Pb amplitudes were more positive, Pb latencies were longer, and Nc amplitudes were less negative in AVdelay trials.

4. Discussion

4.1. Behavior versus EEG activity

The present study investigated neural dynamics of audiovisual temporal fusion processes in 6-month-old infants. As predicted, infants did not show any signs of behavioral discrimination of the 200-ms asynchrony in the habituation-test paradigm. This result confirms earlier work on the temporal window of audiovisual binding in infants in the first year of life (Lewkowicz, 1996).

In contrast, ERP activity patterns differed considerably between the AVsynch and the AVdelay stimuli (Figs. 3–5), indicating perceptual fusion dynamics in the AVdelay trials. In other words, in order to create a simultaneous percept on the behavioral level, the infant brain shows ERP differences in perceptually fused as compared to physically synchronous audiovisual stimuli. This finding may reflect different neural mechanisms involved in synchronous versus fused perception and may be in line with neuroimaging studies demonstrating the involvement of different neural networks in the processing of audiovisually synchronous versus fused stimuli in adult participants (Miller and D'Esposito, 2005; Stevenson et al., 2011). Importantly, ERP modulations elicited by AVdelay as compared to AVsynch stimuli were highly dissimilar from the modulations elicited by stimuli containing a 400-ms audiovisual asynchrony (Kopp and Dietrich, 2013) which were associated with significant asynchrony discrimination on the behavioral level (see detailed discussion below). Moreover, it is important to note that ERP activity patterns of the synchronous condition were highly similar between the previous (Kopp and Dietrich, 2013) and the present

study, indicating a replication of the earlier data and suggesting that synchrony perception did not systematically depend on the presentation of fused or asynchronous stimuli within an experimental session.

4.2. Modulation of auditory-evoked responses

As hypothesized, activity reflected in the auditory ERP components varied with asynchronous visual presentations. In particular, the P2 amplitude was significantly more positive in AVsynch compared to AVdelay stimuli at anterior electrode locations. The P2 amplitude result resembles data in Kopp and Dietrich (2013) however, in that study, the difference between the synchronous and the asynchronous condition did not reach statistical significance. It is conceivable that the neural mechanisms reflected in P2 activity depend on the processing of the auditory event relative to the corresponding visual information, that is, it could be a function of the visual delay. On the other hand, the difference between the two studies could reflect a statistical issue: In the previous study, a smaller sample of participants ($n = 15$) contributed data to the ERP grand averages than in the present experiment ($n = 36$), which could have affected the signal-to-noise ratio. In fact, the absence of a statistically reliable effect of synchronous versus asynchronous trials in Kopp and Dietrich (2013) in the presence of a pronounced P2 amplitude modulation suggest an influence of statistical power. Given the relatively low number of trials infants typically contribute to ERP grand averages and the high interindividual variability of infant EEG data, the reliability of the present results can be regarded as higher, with 36 infants in the final sample (e.g., Talsma and van Harmelen, 2009), however, one has to take into account that the potential sample size effect may limit comparisons between the previous study and the present one.

P2 activity modulations have also been observed in other paradigms investigating audiovisual perception in infants. Hyde et al. (2011) reported larger auditory P2 amplitudes in audiovisually synchronous speech as compared to modified stimuli, including a temporally delayed static face or incongruent dynamic visual information. Therefore, it seems fair to conclude that infants tend to evince greater P2 amplitudes in synchronous audiovisual events than in stimuli with manipulated intersensory temporal relations or congruency. In contrast, in human adults, N1 and P2 amplitudes are typically reported to be diminished in audiovisually synchronous, congruent events as compared to unisensory auditory stimuli, suggesting a processing advantage for multisensory stimuli (e.g., Besle et al., 2004, 2009; Giard and Peronnet, 1999; Stekelenburg and Vroomen, 2007; Van Wassenhove et al., 2005). These amplitude reductions depend on the temporal content and predictability of the visual input and are absent in audiovisual asynchronous stimuli (Pilling, 2009). With respect to the present and previous infant data, it can be concluded that neural activity as reflected in P2 amplitude modulations may be biased toward the perception of synchrony early in perceptual development and may undergo changes until adulthood (see also Hyde et al., 2011).

Critically, the major difference to the auditory-evoked ERP responses in the Kopp and Dietrich (2013) study is

the present lack of latency differences in N1 and P2 as a function of temporal synchrony relations. In the previous experiment, asynchrony perception was associated with significant latency shifts of the auditory ERP components as compared to synchrony, although the auditory event was presented at the same point in time in both conditions. It was hypothesized that a pronounced auditory latency delay may be an indicator for the emergence of an asynchronous percept on the behavioral level. This hypothesis is strengthened by the present data. Perceptual fusion dynamics were associated with ERP amplitude but not latency modulation when compared to synchronous trials.

In adult research, N1 and/or P2 amplitude suppression and latency speed-up in bimodal audiovisual as compared to unisensory auditory stimulus processing (e.g., Pilling, 2009; Stekelenburg and Vroomen, 2007; Van Wassenhove et al., 2005) have been interpreted in terms of facilitatory effects in auditory processing in the bimodal case, an interpretation that is supported by the association with faster reaction times as compared to unisensory processing (e.g., Besle et al., 2004; Giard and Peronnet, 1999). The present and previous (Kopp and Dietrich, 2013) data suggest that simultaneity perception on the behavioral level, which is present in AVsynch and AVdelay stimuli but absent in stimuli with an audiovisual asynchrony of 400 ms, may be associated with faster processing on the neural level as indicated by auditory ERP latencies, at least in infant samples. More research is required into the dynamics of these processes, especially into their development from early stages in ontogeny to adulthood. For example, it would be interesting to follow up by investigating the perception of temporally fused items in older children and adults and relating reaction time data of multisensory perception tasks to the ERP measures. Moreover, this research could help to clarify the nature of perceptual fusion mechanisms in childhood: It has been argued that, before middle childhood, calibration of one sensory modality by another sensory modality may be more likely to occur than real integration (e.g., Burr and Gori, 2012).

4.3. Visual anticipatory activity

N1 and P2 modulations indicate that the hand-clapping sound was perceived depending on the temporal delay of the physical visual input, even when the percept was a simultaneous one. In other words, the preceding visual motion determined subsequent auditory processing. However, the ERP pattern of anticipatory visual processing before the onset of the sound also depended critically on asynchrony or simultaneity perception: Audiovisual asynchrony perception as investigated in Kopp and Dietrich (2013) elicited a polarity reversal of the visual Pb component followed by latency shifts in the Nc component as compared to physically synchronous stimuli. In contrast, here, perceptually fused items were associated with same-polarity amplitude modulations in Pb and Nc when compared with audiovisually synchronous events. That is, neural activity related to predictive mechanisms with respect to the auditory event had a different pattern when

the two modalities were pulled together, resulting in a temporally aligned percept.

The first highly significant ERP modulation was observed in the Pb component. Pb has typically been associated with expectancy processes and mechanisms related to the relevance of presented items (Karrer and Monti, 1995; Kopp and Dietrich, 2013; Kopp and Lindenberger, 2011, 2012; Webb et al., 2005). Here the amplitude of this component was clearly increased and the peak latency was significantly longer in AVdelay than in synchronous trials. It has to be taken into account that the AVdelay items do not contain any fundamentally different visual stimulation but the same biologically possible motion in which only the content is delayed by 200 ms (see Fig. 1). Also, with respect to expectancy processes as to the occurrence of the visual events, it is important to note that AVsynch and AVdelay stimuli were presented randomly with equal probability during the experimental session. Hence, excluding these potential sources of variability, one of the most likely interpretations is that the Pb modulations might reflect mechanisms by which the infants adjust the ongoing neural activity with the to-be-expected auditory activity. As discussed in Kopp and Dietrich (2013), the early ERP patterns – starting approximately 200 ms prior to sound onset – might reflect predictive capacities in the 6-month-olds with respect to temporal synchrony relations. Yet this ERP modulation in fused perception seems substantially different than in behavioral asynchrony perception where activity in the time window of the Pb showed reversed polarity (Kopp and Dietrich, 2013).

There have been only few instances in the literature reporting Pb latency shifts (Kopp and Lindenberger, 2011, 2012; Nikkel and Karrer, 1994; Webb et al., 2005). The present data suggest that variables that affect expectancy processes may result in lengthened or shortened Pb peak latency. It is conceivable that longer neural processing of AVdelay stimuli can be associated with a temporal shift of activity toward fusion with the expected hand-clapping sound. Following this line of argument, one could assume that the peak latency difference of approximately 20 ms between AVdelay and AVsynch trials is likely to reflect a temporal interaction process between the visual and the anticipated auditory modality, as this difference is significant on the one hand but much smaller than the 200-ms visual delay on the other hand. Alternatively, one could regard the Pb latency in AVdelay trials as a consequence of the pronounced Pb amplitude increase, thus resulting in the later peak of the more positive ERP deflection.

Here, Nc amplitudes were significantly less negative for AVdelay than for synchronous stimuli. Activity related to the infant Nc component has been associated with top-down, attentional mechanisms and with memory processes (e.g., Ackles, 2008; Richards, 2003). In a number of studies, Nc amplitudes have been observed to decrease in familiar as compared to unfamiliar items (e.g., Kopp and Lindenberger, 2011; Reynolds and Richards, 2005) and to change with repeated stimulus presentation (Nikkel and Karrer, 1994; Stets and Reid, 2011; Wiebe et al., 2006). In line with this research, Hyde et al. (2011) interpreted the Nc amplitude modulations in multisensory processing observed in infants in terms of an attentional and novelty

account. They concluded that infants found altered audiovisual stimuli more interesting and more novel, which caused them to attend to them more than to the congruent, synchronous stimuli.

The Nc pattern in the present study, which revealed a pronounced amplitude decrease in AVdelay compared to AVsynch trials, is not clear. First, temporal fusion of the sensory input as an adjustment of visual and anticipated auditory activity might correlate with this very specific Nc modulation. Following the attentional or novelty interpretation of Nc, however, one would then raise the question why AVdelay stimuli should be perceived as less interesting or more familiar. Second, based on the behavioral data, the Nc amplitude modulation would be difficult to explain in terms of familiarity versus novelty, given that both AVdelay and AVsynch stimuli seemed to create simultaneous percepts in this sample of 6-month-olds. Hence, the two stimuli should be perceived as equally familiar. Third, an alternative, viable explanation might be that the Nc pattern in AVdelay trials could result from a possible carry-over effect of the preceding Pb amplitude response. Earlier studies have shown that Pb and Nc activity presumably reflect independent or at least not fully overlapping neural mechanisms (e.g., Kopp and Lindenberger, 2011; Nikkel and Karrer, 1994; Webb et al., 2005). However, a paradigm like the present one has not been tested before and, thus, it is largely unknown whether and/or how Pb and Nc are dissociable in this context.

In sum, Pb and Nc responses suggest that prior to sound onset, neural activity showed pronounced modulations depending on how temporal synchrony relations would be perceived behaviorally, and that this predictive listening followed expectancy processes. The work reported here mirrors not only results from multisensory research in adult participants (e.g., Stekelenburg and Vroomen, 2007; Van Wassenhove et al., 2005; Vroomen and Stekelenburg, 2009) but also experimental findings on the function of internal predictive mechanisms in the adult auditory system (e.g., Pieszek et al., 2013; Timm et al., 2013).

As in the Kopp and Dietrich (2013) study, differences between the conditions were observed primarily at anterior electrode sites. The topographic pattern suggests that similar neural networks may be involved in asynchrony and fused perception. Nevertheless, the timing and the dynamics of the ERP deflections seem very different between the two types of stimuli. In adults, neuroimaging data have shown differential brain activation in audiovisual synchrony and asynchrony versus fused perception (e.g., Miller and D'Esposito, 2005; Stevenson et al., 2011), and the ERP curves here also suggest differential activation in the infant sample. In future studies, source localization or related methods could disentangle the different contributions of the neural networks involved. Furthermore, additional control conditions may help identify contributions from the unisensory stimulus components when aiming for additive modeling. For a very detailed discussion of the pros and cons of the present experimental paradigm, possible contributing factors, and specific considerations about testing EEG in infants using this or similar paradigms, the reader is referred to the article by Kopp and Dietrich (2013).

To conclude, this study investigated audiovisual temporal fusion mechanisms in 6-month-old infants and found pronounced differences in ERP activity between synchronous and perceptually fused trials while behavioral data indicated simultaneity perception in both stimuli. The auditory-evoked ERP modulations suggest that the sound of the multisensory stimulus is processed differently when the magnitude of the audiovisual temporal disparity allows for perceptual fusion than when the resulting percept is asynchronous. The preceding visual motion had a strong influence on neural processing prior to sound onset, suggesting anticipatory mechanisms in multisensory temporal perception. Indeed, anecdotal evidence shows that children even younger than six months can be observed anticipating the sound of a toy hitting the ground and being startled before the sound actually occurs. Together with this and many other observations, the present experimental findings suggest remarkable predictive capacities as to audiovisual temporal synchrony relations very early in ontogeny.

Conflict of interest statement

Any actual or potential conflict of interest is disclosed including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.dcn.2014.01.001.

References

Ackles, P.K., 2008. Stimulus novelty and cognitive-related ERP components of the infant brain. Percept. Mot. Skills 106, 3–20.

Ackles, P.K., Cook, K.G., 2007. Attention or memory? Effects of familiarity and novelty on the Nc component of event-related brain potentials in six-month-old infants. Int. J. Neurosci. 117, 837–867.

Aslin, 2007. What's in a look? Dev. Sci. 10, 48–53.

Bahrick, L.E., 1983. Infants' perception of substance and temporal synchrony in multimodal events. Infant Behav. Dev. 6, 429–451.

Bahrick, L.E., 1987. Infants' intermodal perception of two levels of temporal structure in natural events. Infant Behav. Dev. 10, 387–416.

Besle, J., Bertrand, O., Giard, M.-H., 2009. Electrophysiological (EEG, sEEG, MEG) evidence for multiple audiovisual interactions in the human auditory cortex. Hear. Res. 258, 143–151.

Besle, J., Fort, A., Delpuech, C., Giard, M.H., 2004. Bimodal speech: early suppressive visual effects in the human auditory cortex. Eur. J. Neurosci. 20, 2225–2234.

Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., Mangin, J.-F., 2009. Hearing faces: how the infant brain matches the face it sees with the speech it hears. J. Cogn. Neurosci. 21, 905–921.

Burr, D., Gori, M., 2012. Multisensory integration develops late in humans. In: Murray, M.M., Wallace, M.T. (Eds.), The Neural Basis of Multisensory Processes. CRC Press, Boca Raton (FL).

Bushara, K.O., Grafman, J., Hallett, M., 2001. Neural correlates of auditory–visual stimulus onset asynchrony detection. J. Neurosci. 21, 300–304.

Colombo, J., Mitchell, D.W., 2009. Infant visual habituation. Neurobiol. Learn. Mem. 92, 225–234.

Dhamala, M., Assisi, C.G., Jirsa, V.K., Steinberg, F.L., Kelso, J.A.S., 2007. Multisensory integration for timing engages different brain networks. Neuroimage 34, 764–773.

Dixon, N.F., Spitz, L., 1980. The detection of auditory visual desynchrony. Perception 9, 719–721.

Dodd, B., 1979. Lip reading in infants: attention to speech presented in- and out-of-synchrony. Cogn. Psychol. 11, 478–484.

Fendrich, R., Corballis, P.M., 2001. The temporal cross-capture of audition and vision. Percept. Psychophys. 63, 719–725.

Fujisaki, W., Shimojo, S., Kashino, M., Nishida, S., 2004. Recalibration of audiovisual simultaneity. Nat. Neurosci. 7, 773–778.

Giard, M.H., Peronnet, F., 1999. Auditory–visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J. Cogn. Neurosci. 11, 473–490.

Hillock, A.R., Powers, A.R., Wallace, M.T., 2011. Binding of sights and sounds: age-related changes in multisensory temporal processing. Neuropsychologia 49, 461–467.

Hollich, G., Newman, R.S., Juszyk, P.W., 2005. Infants' use of synchronized visual information to separate streams of speech. Child Dev. 76, 598–613.

Hyde, D.C., Jones, B.L., Flom, R., Porter, C.L., 2011. Neural signatures of face–voice synchrony in 5-month-old human infants. Dev. Psychobiol. 53, 359–370.

Hyde, D.C., Jones, B.L., Porter, C.L., Flom, R., 2010. Visual stimulation enhances auditory processing in 3-month-old infants and adults. Dev. Psychobiol. 52, 181–189.

Karrer, R., Monti, L.A., 1995. Event-related potentials of 4–6-week-old infants in a visual recognition memory task. Electroencephalogr. Clin. Neurophysiol. 94, 414–434.

King, A.J., 2005. Multisensory integration: strategies for synchronization. Curr. Biol. 15, R339–R341.

Kopp, F., Dietrich, C., 2013. Neural dynamics of audiovisual synchrony and asynchrony perception in 6-month-old infants. Front. Psychol. 4, 2, http://dx.doi.org/10.3389/fpsyg.2013.00002.

Kopp, F., Lindenberger, U., 2011. Effects of joint attention on long-term memory in 9-month-old infants: an event-related potentials study. Dev. Sci. 14, 660–672.

Kopp, F., Lindenberger, U., 2012. Social cues at encoding affect memory in four-month-old infants. Soc. Neurosci. 7, 458–472.

Lewkowicz, D.J., 1986. Developmental changes in infants' bisensory response to synchronous durations. Infant Behav. Dev. 9, 335–353.

Lewkowicz, D.J., 1992. Infants' response to temporally based intersensory equivalence: the effect of synchronous sounds on visual preferences for moving stimuli. Infant Behav. Dev. 15, 297–324.

Lewkowicz, D.J., 1996. Perception of auditory–visual temporal synchrony in human infants. J. Exp. Psychol. Hum. Percept. Perform. 22, 1094–1106.

Lewkowicz, D.J., 2000. The development of intersensory temporal perception: an epigenetic systems/limitations view. Psychol. Bull. 126, 281–308.

Lewkowicz, D.J., 2010. Infant perception of audiovisual speech synchrony. Dev. Psychol. 46, 66–77.

Lewkowicz, D.J., Leo, I., Simion, F., 2010. Intersensory perception at birth: newborns match nonhuman primate faces and voices. Infancy 15, 46–60.

Lewkowicz, D.J., Sowinski, R., Place, S., 2008. The decline of cross-species intersensory perception in human infants: underlying mechanisms and its developmental persistence. Brain Res. 1242, 291–302.

Macaluso, E., George, N., Dolan, R., Spence, C., Driver, J., 2004. Spatial and temporal factors during processing of audiovisual speech: a PET study. Neuroimage 21, 725–732.

Miller, L.M., D'Esposito, M., 2005. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. J. Neurosci. 25, 5884–5893.

Navarra, J., Alsius, A., Velasco, I., Soto-Faraco, S., Spence, C., 2010. Perception of audiovisual speech synchrony for native and non-native language. Brain Res. 1323, 84–93.

Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., Spence, C., 2005. Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. Cogn. Brain Res. 25, 499–507.

Nikkel, L., Karrer, R., 1994. Differential effects of experience on the ERP and behavior of 6-month-old infants. Dev. Neuropsychol. 10, 1–11.

Pilling, M., 2009. Auditory event-related potentials (ERPs) in audiovisual speech perception. Journal of Speech Language, and Hearing Research 52, 1073–1081.

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1990. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187.

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C.H., Avanzini, F., Puce, A., Pollick, F.E., 2009. Multisensory integration of drumming actions: musical expertise affects perceived audiovisual synchrony. Exp. Brain Res. 198, 339–352.

Pieszek, M., Widmann, A., Gruber, T., Schröger, E., 2013. The human brain maintains contradictory and redundant auditory sensory predictions. PLoS ONE 8, e53634, http://dx.doi.org/10.1371/journal.pone.0053634.

Powers, A.R., Hillock, A.R., Wallace, M.T., 2009. Perceptual training narrows the temporal window of multisensory binding. J. Neurosci. 29, 12265–12274.

Reynolds, G.D., Richards, J.E., 2005. Familiarization, attention, and recognition memory in infancy: an event-related potential and source-localization study. Dev. Psychol. 4, 598–615.

Richards, J.E., 2003. Attention affects the recognition of briefly presented visual stimuli in infants: an ERP study. Dev. Sci. 6, 312–328.

Spelke, E.S., 1979. Perceiving bimodally specified events in infancy. Dev. Psychol. 15, 626–636.

Spence, C., Squire, S., 2003. Multisensory integration: maintaining the perception of synchrony. Curr. Biol. 13, R519–R521.

Stekelenburg, J.J., Vroomen, J., 2007. Neural correlates of multisensory integration of ecologically valid audiovisual events. J. Cogn. Neurosci. 19, 1964–1973.

Stets, M., Reid, V., 2011. Infant ERP amplitudes change over the course of an experimental session: implications for cognitive processes and methodology. Brain Dev. 33, 558–568.

Stevenson, R.A., Altieri, N.A., Kim, S., Pisoni, D.B., James, T.W., 2010. Neural processing of asynchronous audiovisual speech perception. Neuroimage 49, 3308–3318.

Stevenson, R.A., VanDerKlok, R.M., Pisoni, D.B., James, T.W., 2011. Discrete neural substrates underlie complementary audiovisual speech integration processes. Neuroimage 55, 1339–1345.

Stevenson, R.A., Zemtsov, R.K., Wallace, M.T., 2012. Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. J. Exp. Psychol. Hum. Percept. Perform. 38, 1517–1529.

Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay between attention and multisensory integration. Trends Cogn. Sci. 14, 400–410.

Talsma, D., Senkowski, D., Woldorff, M.G., 2009. Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. Exp. Brain Res. 198, 313–328.

Talsma, D., van Harmelen, A.-L., 2009. Procedures and strategies for optimizing the signal-to-noise ratio in event-related potential data. In: Handy, T.C. (Ed.), Brain Signal Analysis: Advances in Neuroelectric and Neuromagnetic Methods. The MIT Press, Cambridge (Mass.).

Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neurosci. 14, 2, http://dx.doi.org/10.1186/1471-2202-14-2.

Van Wassenhove, V., Grant, K.W., Poeppel, D., 2005. Visual speech speeds up the neural processing of auditory speech. PNAS 102, 1181–1186.

Van Wassenhove, V., Grant, K.W., Poeppel, D., 2007. Temporal window of integration in auditory–visual speech perception. Neuropsychologia 45, 598–607.

Vatakis, A., Navarra, J., Soto-Faraco, S., Spence, C., 2007. Temporal recalibration during asynchronous audiovisual speech perception. Exp. Brain Res. 181, 173–181.

Vatakis, A., Spence, C., 2006. Audiovisual synchrony perception for music, speech, and object actions. Brain Res. 1111, 134–142.

Vroomen, J., Stekelenburg, J.J., 2009. Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. J. Cogn. Neurosci. 22, 1583–1596.

Webb, S.J., Long, J.D., Nelson, C.A., 2005. A longitudinal investigation of visual event-related potentials in the first year of life. Dev. Sci. 8, 605–616.

Wiebe, S.A., Cheatham, C.L., Lukowski, A.F., Haight, J.C., Muehleck, A.J., Bauer, P.J., 2006. Infants' ERP responses to novel and familiar stimuli change over time: implications for novelty detection and memory. Infancy 9, 21–44.