



Published in final edited form as:

*Dev Psychobiol.* 2014 April ; 56(3): 355–372. doi:10.1002/dev.21104.

## Neural Correlates of Intersensory Processing in Five-Month-Old Infants

Greg D. Reynolds<sup>1</sup>, Lorraine E. Bahrick<sup>2</sup>, Robert Lickliter<sup>2</sup>, and Maggie W. Guy<sup>1</sup>

<sup>1</sup>University of Tennessee Knoxville

<sup>2</sup>Florida International University

### Abstract

Two experiments assessing event-related potentials in 5-month-old infants were conducted to examine neural correlates of attentional salience and efficiency of processing of a visual event (woman speaking) paired with redundant (synchronous) speech, nonredundant (asynchronous) speech, or no speech. In Experiment 1, the Nc component associated with attentional salience was greater in amplitude following synchronous audiovisual as compared with asynchronous audiovisual and unimodal visual presentations. A block design was utilized in Experiment 2 to examine efficiency of processing of a visual event. Only infants exposed to synchronous audiovisual speech demonstrated a significant reduction in amplitude of the late slow wave associated with successful stimulus processing and recognition memory from early to late blocks of trials. These findings indicate that events that provide intersensory redundancy are associated with enhanced neural responsiveness indicative of greater attentional salience and more efficient stimulus processing as compared with the same events when they provide no intersensory redundancy in 5-month-old infants.

### Keywords

intersensory perception; event-related potentials; infancy; attention

### Introduction

It is well established that multimodal stimulation is highly salient and promotes heightened attention, perceptual processing and memory in human infants and adults as well as nonhuman animal infants (Bahrick & Lickliter, 2000; 2002; 2012; Lewkowicz, 2000; Lickliter & Bahrick, 2000). What accounts for the attentional salience of multimodal stimulation? Multimodal stimulation often provides intersensory redundancy, the synchronous co-occurrence of the same amodal information (e.g., rhythm, tempo, intensity changes) across two or more sense modalities (Bahrick & Lickliter, 2000). Infants are routinely exposed to redundant (e.g., synchronous faces and voices) and nonredundant (e.g., faces or voices alone, or one person's face moving out of synchrony with another person's voice) stimulation in their everyday lives.

Bahrick and Lickliter (2000, 2002) have proposed the intersensory redundancy hypothesis (IRH), a model that describes the central role of selective attention to intersensory redundancy (i.e., temporal synchrony) in guiding early perceptual and cognitive development. Research has demonstrated that selective attention to intersensory redundancy

is a cornerstone of perceptual learning and early cognitive development (Bahrick & Lickliter, 2000, 2012; Lewkowicz, 2000). The IRH proposes that temporal synchrony across two or more sensory systems promotes attention to redundantly specified properties of objects and events (e.g., rhythm, tempo, affect) at the expense of other nonredundantly specified properties, particularly in early development when attentional resources are most limited.

Findings at the behavioral level have provided support for the intersensory redundancy hypothesis. For example, Bahrick and Lickliter (2000) habituated 5-month-old infants to a hammer tapping a complex rhythm in redundant (synchronous audiovisual) versus nonredundant (unimodal visual, unimodal auditory, or asynchronous audiovisual) stimulation. Only infants habituated to the rhythm in the synchronous condition demonstrated dishabituation (i.e., increased looking) to a change in rhythm. Similarly, Flom & Bahrick (2007) found that infants were able to discriminate the affect of a woman speaking by 4 months of age when exposed to redundant stimulation (synchronous audiovisual speech) but not nonredundant stimulation (unimodal visual, unimodal auditory or asynchronous audiovisual speech). Discrimination of affect during unimodal visual and unimodal auditory speech emerged later in development. Thus, the redundant, synchronous presentation of affective information across the auditory and visual sense modalities enhanced infants' perceptual processing of amodal information, a process referred to as intersensory facilitation (Bahrick & Lickliter, 2002, 2012). The consistent finding (Bahrick, Flom, & Lickliter, 2002; Bahrick & Lickliter, 2000, 2012; Flom & Bahrick, 2007) that young infants are able to abstract amodal stimulus properties of redundant audiovisual stimuli at an earlier age than the same amodal stimulus properties can be abstracted from nonredundant audiovisual stimuli indicates that intersensory redundancy fosters enhanced perceptual processing in early infancy.

What are the underlying mechanisms that support enhanced perceptual processing of redundant information in early development? To date most developmental research has focused on the behavioral level. Behavioral findings indicate that infants are able to perceive amodal information provided by multimodal sensory stimulation at a very early age (for reviews see Bahrick & Pickens, 1994; Lewkowicz, 2000; Lickliter & Bahrick, 2000; Walker-Andrews, 1997). Research suggests that this ability guides young infants' selective attention and is fundamental to their unitary perception of meaningful events (see Bahrick & Lickliter, 2002; 2012 Gibson & Pick, 2000). Although a large body of research demonstrates impressive intersensory processing skills in human and nonhuman animal infants (Bremner, Lewkowicz, & Spence, 2012; Calvert, Spence, & Stein, 2004; Lewkowicz & Lickliter, 1994), there is currently little understanding of potential mechanisms underlying this skill/process. For example, it is not known if redundantly presented information is processed more efficiently because intersensory redundancy serves as a salient, attention-getting stimulus (Cohen, 1972) or if redundant information is just easier to process, but no more salient, than nonredundant information. There is also little known about the neural processes involved in processing intersensory redundancy in infancy.

The majority of what we do know about neural processes involved in multimodal perception in early development is based on comparative work with non-human animal subjects (e.g., Jay & Sparks, 1984; Stein, Meredith, & Wallace, 1994; Wallace & Stein, 1997; Wallace, Wilkinson, & Stein, 1996). This line of research has demonstrated "superadditive" effects of multimodal stimulation on firing rates of neurons in the superior colliculus of young cats (Wallace & Stein, 1997) and monkeys (Jay & Sparks, 1984; Wallace, Wilkinson, & Stein, 1996). Furthermore, comparative work with cats and primates indicates that multisensory neurons are distributed throughout the cerebral cortex, including areas classically viewed as unisensory domains (for reviews see Ghazanfar & Schroeder, 2006; Stein & Stanford,

2008). Cortical areas commonly identified as multisensory regions in monkeys include the superior temporal sulcus (STS; Bruce, Desimone, & Gross, 1981; Hikosaka, 1993), the intraparietal complex (IP; Mazzone et al., 1999; Linden et al., 1999), and the frontal cortex (Benevento et al., 1977).

Neuroimaging studies using human participants have also demonstrated that these areas are involved in multimodal processing in adulthood (e.g., Beauchamp, 2005; Gobbelè et al., 2003; Lutkenhoner et al., 2002). For example, research indicates the STS is actively involved in processing audiovisual speech in adult participants (Senkowski, Saint-Amour, Gruber, & Foxe, 2008), and recent fMRI work (Marchant, Ruff, & Driver, 2012) demonstrates a significantly higher BOLD response to synchronous audiovisual stimuli compared to asynchronous audiovisual stimuli in the: STS, superior temporal gyrus, thalamus, and putamen. Furthermore, Laurienti and colleagues (2003) found increased BOLD responses occurring in the anterior cingulate gyrus and medial prefrontal cortex to matching (or congruent) audiovisual stimuli compared to non-matching (incongruent) audiovisual stimuli. Interestingly, studies utilizing cortical source analyses with infant participants indicate that these frontal areas are likely sources of the Nc ERP component associated with infant attention and visual preferences (Reynolds, Courage, & Richards, 2010; Reynolds & Richards, 2005). Thus, research across species demonstrates that multiple cortical and subcortical areas are involved in multimodal processing; however, little is known about neural processing of multimodal stimulation in infancy, due in part to practical and ethical concerns related to use of standard neuroimaging techniques (e.g., PET, fMRI) with human infants (Reynolds & Richards, 2009). Here we focus on one aspect of neural processing of multimodal stimulation in infancy, the neural underpinnings of attention to intersensory redundancy.

Although behavioral (i.e., habituation) findings indicate that intersensory redundancy promotes selective attention to and perceptual processing of amodal stimulus properties in infancy, the underlying neural processes are relatively unknown and the point in the information processing stream at which facilitation occurs cannot be determined based on behavioral findings alone. The ERP is a particularly useful measure for examining component processes (e.g., orienting, attention, memory) involved in perceptual and cognitive processing that potentially occur within the course of a single look (see, Reynolds & Guy, 2012). The ERP represents voltage oscillations in the electroencephalogram (EEG) that are time-locked with a specific event of interest. The ERP is averaged across trials to increase the signal-to-noise ratio in the EEG, and components can be identified in the averaged waveform that are associated with specific aspects of perceptual and cognitive processing. ERP work with adult and infant participants indicates that components associated with early auditory and visual processing are significantly greater in amplitude following multimodal audiovisual stimulus presentations when compared to the sum of unimodal auditory and visual presentations (e.g., Giard & Peronnet, 1999; Hyde, Jones, Porter, & Flom, 2010; Molholm et al., 2002; Santangelo, Van der Lubbe, Olivetti-Berlardinelli, & Postma, 2008).

In infant ERP research, the Nc component has been shown to be associated with visual attention and stimulus salience (Courchesne, Ganz, & Norcia, 1981; de Haan & Nelson, 1997; Reynolds et al., 2005, 2010; Richards, 2003). The Nc is a negatively-polarized ERP component that occurs between 350 and 750 ms after stimulus onset over midline electrodes. A common finding across early infant ERP studies was that Nc is greater in amplitude following oddball (or rare) stimulus presentations than following standard stimulus presentations (Courchesne, 1977; Courchesne, Ganz, & Norcia, 1981; Karrer & Ackles, 1987, 1988; Karrer & Monti, 1995; Nikkel & Karrer, 1994). More recent findings indicate that Nc amplitude is impacted by stimulus salience as opposed to frequency of

presentation or novelty per se (de Haan & Nelson, 1997, 1999; Reynolds & Richards, 2005; Reynolds, Courage, and Richards, 2010). Reynolds, Courage, and Richards (2010) integrated a behavioral measure of infant visual preferences (i.e., paired comparison trials) into an ERP study and found that Nc is greatest in amplitude to the infant's preferred stimulus regardless of novelty or familiarity. These findings indicate that the Nc component is associated with infant visual attention and varies in amplitude based on stimulus salience (Courchesne et al., 1981; Nelson, 1994; Reynolds et al., 2010; Richards, 2003). Thus, if intersensory redundancy recruits infant attention, then Nc would be expected to be greater in amplitude to redundant multimodal stimulation than nonredundant stimulation.

The late slow wave (LSW) is believed to reflect stimulus encoding, and differential amplitude of the LSW based on stimulus type is an electrophysiological index of infant recognition memory (see review, de Haan, 2007). The LSW typically occurs from 1 to 2 s following stimulus onset over temporal and frontal leads. A consistent finding across studies is that the LSW demonstrates a reduction in amplitude with repeated stimulus presentations that is indicative of stimulus encoding (de Haan & Nelson, 1999; Reynolds, Guy, & Zhang, 2011; Snyder, 2010; Snyder, Webb, & Nelson, 2002; Snyder et al., 2010; Webb, Long, & Nelson, 2005; Wiebe et al., 2006). Given the behavioral findings demonstrating enhanced processing of redundant multimodal stimuli in infancy (Bahrick, Flom, & Lickliter, 2002; Bahrick & Lickliter 2000; Flom & Bahrick, 2007), infants would be expected to require less exposure to a redundant multimodal stimulus in comparison to nonredundant stimulus (multimodal or unimodal) in order to demonstrate a significant reduction in LSW amplitude. Such an effect would provide evidence of more efficient processing of redundant stimuli at the neural level.

In this paper, we describe two studies designed to test neural mechanisms underlying the salience and enhanced processing of intersensory redundancy. For consistency with prior behavioral research in this area (e.g., Flom and Bahrick, 2007), we exposed 5-month-old infants to videos of a woman speaking providing intersensory redundancy (synchronous audiovisual speech) versus no redundancy (asynchronous audiovisual or unimodal visual speech). Five-month-old infants are skilled at discriminating synchrony from asynchrony and at detecting amodal properties, including rhythm, tempo, and affective information common to faces and voices (Bahrick & Lickliter, 2012). Experiment 1 was designed to test the attentional salience of intersensory redundancy as reflected by the amplitude of the Nc component. If intersensory redundancy provided by multimodal stimulation is highly salient and captures infant attention (consistent with the IRH, Bahrick & Lickliter, 2000), then infants should show a greater amplitude Nc component to stimuli depicting intersensory redundancy (synchronous audiovisual) than to stimuli depicting no intersensory redundancy (asynchronous audiovisual, or unimodal visual). Experiment 2 was designed to learn more about mechanisms underlying intersensory processing. If intersensory redundancy promotes enhanced perceptual processing in comparison to nonredundant stimulation, then this should be reflected by significant changes in the LSW with repeated exposure to a redundant multimodal stimulus. If infants demonstrate enhanced intersensory processing due to deeper levels of attentional engagement, then changes in LSW amplitude over time should be paired with greater amplitude Nc to redundant audiovisual stimuli.

## **Experiment 1: Intersensory Redundancy and Attentional Salience as assessed by the Nc Component**

In Experiment 1, we utilized high-density EEG to examine the impact of multimodal (synchronous and asynchronous) audiovisual and unimodal visual stimulus presentations on ERP components associated with visual attention and face processing. We tested infants at 5 months of age for consistency with previous behavioral work in the area (e.g., Bahrick &

Lickliter, 2000). We exposed infants to repeated presentations of a woman speaking a short phrase under conditions depicting three stimulus types: unimodal visual (video of a woman speaking with no soundtrack), synchronous audiovisual (video of a woman speaking with synchronous soundtrack), and asynchronous audiovisual (video a woman speaking with temporally asynchronous soundtrack).

Three studies (Grossmann, Striano, & Friederici, 2006; Hyde, Jones, Flom, & Porter, 2011; Vogel, Monesson, & Scott, 2012) have examined the effects of audiovisual face-voice pairings on the Nc component in infant participants. Findings from these studies have been somewhat inconsistent regarding the effects of congruency across auditory and visual stimulus components. For example, Grossmann and colleagues (2006) found that infants demonstrate greater negativity to face-voice pairings conveying incongruent emotional information compared to face-voice pairings conveying congruent emotional information. In contrast, Vogel and colleagues (2012) found that infants demonstrate greater amplitude Nc, indicating greater attention, to face-voice pairings conveying congruent emotional information.

Vogel and colleagues (2012) speculated that inconsistency across studies in the direction of congruency effects may be due to differences in stimuli used or task design. Grossmann and colleagues (2006) first presented a face displaying a happy or angry facial expression and then presented an audio clip of a woman speaking in a happy or angry tone. The ERPs were time-locked to the audio presentation of the woman speaking. In contrast, Vogel and colleagues first played an audio clip of a woman speaking in a happy or sad tone, and then presented a face displaying a happy or sad facial expression. The ERPs were time-locked to the visual presentation of the face, which is typical of most studies examining Nc. Thus, the inconsistency may have been due to the fact that in one study the audio component always served as the source of information for detecting congruity or incongruity (Grossmann et al., 2006), and in the other study the video component always served as the source of information for detecting congruity or incongruity (Vogel et al., 2012). It also worthwhile to note that the “negative component” analyzed in response to the audio clips in the Grossman and colleagues’ study (2006) was actually a relatively slight negative deflection that occurred following a high amplitude positive-going change in the ERP waveform. Thus, their “negative component” was actually positively-polarized, and the “greater negativity” observed in response to incongruent stimuli could just as easily be interpreted as a greater amplitude positive component occurring in response to congruent stimuli. Since the audio and video components of the face-voice pairings were not presented in synchrony in either of these studies, the stimuli used did not provide intersensory redundancy.

In the only published study to date to specifically examine the effects of redundancy provided through audiovisual synchrony on infant ERPs, Hyde and colleagues (2011) exposed 5-month-old infants to synchronous and asynchronous audiovisual presentations of a woman speaking. In the synchronous condition, infants heard an audio clip paired simultaneously with a matching video clip (i.e., video and audio of a woman saying, “Hi baby”). In the asynchronous condition, the infants heard the same audio clip paired with the simultaneous presentation of a non-matching video clip (i.e., video clip showed a woman mouthing, “You’re such a beautiful baby”). In contrast to what would be expected based on the behavioral literature (Bahrck & Lickliter, 2000; Flom & Bahrck, 2007), the authors found that infants demonstrated greater amplitude Nc to the asynchronous face-voice pairings compared to the synchronous face-voice pairings, and concluded that the greater amplitude Nc to asynchronous audiovisual stimuli reflected detection of a novel stimulus category.

Although their conclusion regarding this effect may be correct, Hyde and colleagues did not fully balance the audio and visual components in their synchronous and asynchronous conditions. The phrase used as the audio component, “Hi baby,” remained constant across all stimulus presentations. Thus, the video clip of a woman mouthing, “You’re such a beautiful baby,” only occurred in the asynchronous condition and was incongruent with the ongoing, repeated auditory presentations of the phrase, “Hi baby.” Given that Nc is greater in amplitude to low-frequency or oddball stimuli (Courchesne et al., 1977, 1981; Reynolds & Richards, 2005) and the amplitude of Nc is likely influenced by overall procedural context (Richards, 2003), the presentation of the non-matching video clip may have led to an oddball effect occurring against the standard presentation of the auditory stimulus (i.e., the “Hi baby” phrase) in the Hyde and colleagues’ (2011) study.

We used a balanced design for our synchronous and asynchronous stimulus conditions in the current study, utilizing two different phrases for both the video and audio components of our stimuli to avoid creating a “standard” stimulus and potential “oddball” effects. Consistent with the intersensory redundancy hypothesis and behavioral findings (Bahrick & Lickliter, 2000, 2002, 2012), we predicted that with the greater level of control in the current study, infants would show greater attention to synchronous audiovisual presentations and this would be associated with greater amplitude Nc when compared to asynchronous audiovisual and unimodal visual trials. Greater amplitude Nc in the synchronous compared to asynchronous condition would allow us to rule out the possibility that the differences across groups were simply based on additive effects (audio plus visual as compared with visual only). Amount and type (auditory and visual) of stimulation were equated across synchronous and asynchronous conditions and only the redundancy differed between them.

A secondary goal of Experiment 1 was to conduct a more exploratory and descriptive analysis of the potential impact of intersensory redundancy on ERP components associated with face processing and speech perception in infancy. The ERP components of interest for these analyses included the N290 and P400 components associated with face processing (e.g., de Haan, Johnson, & Halit, 2007; Farroni, Csibra, Simion, Johnson, 2002; Halit, de Haan, & Johnson, 2003); and the auditory P1 and N250 components associated with speech perception (Benasich et al., 2006; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Rivera-Gaxiola, Silva-Pereya, & Kuhl, 2005).

The N290 and P400 have been identified as ERP components related to face-processing in infancy (de Haan, Johnson, & Halit, 2007). The N290 is a negatively-polarized component that occurs over midline and posterior electrodes with peak latency between 290 and 350 ms after stimulus onset (Halit, de Haan, & Johnson, 2003). By 3 months of age, the N290 is greater in amplitude to faces than noise (Halit, Csibra, Volein, & Johnson, 2004). The P400 is a positive-going component that occurs over posterior midline and lateral electrodes and reaches peak amplitude between 390 and 450 ms after stimulus onset. By 6 months of age, the P400 has a shorter latency to peak in response to faces than objects (de Haan & Nelson, 1999), and by 12 months of age, the P400 is shorter in latency to upright versus inverted human faces (Halit et al., 2003).

The auditory P1 component is the first positive peak (also referred to as P150) in the ERP waveform that occurs across the scalp with a peak latency between 150 – 250 ms after stimulus onset. The P1 is sensitive to native and non-native speech contrasts for 6 and 12 month olds (Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Rivera-Gaxiola, Silva-Pereya, & Kuhl, 2005), and is similar in latency to the auditory P2 component that has been shown to be associated with auditory recognition memory in newborn infants (e.g., de Regnier, Nelson, Thomas, Wewerka, & Georgieff, 2000; de Regnier, Wewerka, Georgieff, Mattia, & Nelson, 2002). ERP studies on audiovisual speech perception in infants have had

inconsistent results. For example, some studies have found a reduction in amplitude of early auditory components in response to phonemes preceded by (Bristow et al., 2008) or paired with (Kushnerenko, Teinonen, Volein, & Csibra, 2008) congruent visual cues (van Wassenhove, Grant, & Poeppel, 2005). In contrast, Hyde and colleagues' (2011) found increased amplitude of early auditory components in response to speech paired with congruent visual cues compared to speech paired with incongruent visual cues.

With respect to our secondary analyses of face processing components (N290 and P400) and speech processing components (auditory P1), we made no specific predictions regarding differential effects of redundant and nonredundant audiovisual stimuli. However, due to the potential additive effects of combining auditory and visual stimulation, we predicted that both audiovisual conditions (synchronous and asynchronous) would be associated with greater amplitude ERP across these components compared to the unimodal visual condition. Because the auditory and visual components of the asynchronous stimulus we used were spatially co-located and contained synchronous stimulus onset, we reasoned that basic multimodal additive effects would occur in both audiovisual conditions, but the predicted attention-related effect of intersensory redundancy on Nc amplitude would only occur in the synchronous audiovisual condition.

## Method

**Participants**—A sample of 15 infants (9 male, 6 female) was tested at 5 months of age. Infants were tested within a week of their 22 week birthdate. Only infants born full term (at least 38 weeks gestation) without complications and of normal birth weight were recruited. Participants were drawn from a predominantly Caucasian and middle-class population. The ethnic/racial distribution of participants was: 14 Caucasian (not Hispanic), and 1 Biracial. An additional 23 infants were tested, but not included in the final sample due to fussiness/distractibility ( $N = 6$ ), excessive artifact in the EEG ( $N = 13$ ), and technical problems ( $N = 4$ ). This level of attrition falls within the typical range of 50 – 75 % for infant ERP studies (DeBoer, Scott, & Nelson, 2007).

**Apparatus**—Participants were positioned on their parent's lap in a sound-attenuated room. They were seated 55 cm away from a 27" color LCD monitor (Dell 2707 WFP) with 60 Hz resolution. Speakers were positioned directly behind the monitor for presenting the auditory components of bimodal stimulus presentations. A digital camcorder (Sony DCR-HC28) was located just below the monitor in order to judge infant visual fixations. Fixations were judged online using a video feed to a computer in the experiment control room, adjacent to the testing room. The video was recorded through use of Netstation software produced by Electrical Geodesics Incorporated (EGI). The Netstation was used to record EEG data and to synchronize this data with the video.

## Stimuli

**Test Stimuli:** Infants were exposed to three different stimulus types: unimodal visual, synchronous audiovisual, and asynchronous audiovisual speech. Importantly, two exemplars (depicting different phrases) were used for each stimulus type resulting in a total of 6 test stimuli. The unimodal visual stimuli consisted of dynamic videos without soundtracks, the synchronous audiovisual stimuli consisted of dynamic videos with temporally matching soundtracks, and the asynchronous audiovisual stimuli consisted of dynamic videos with temporally mis-matching soundtracks. All three stimulus types consisted of a female adult actress reciting one of two phrases ("Come over here by me!" or "Where's the baby going?") in infant-directed speech using positive affect. For the asynchronous audiovisual condition, the soundtracks were reversed. For example, the video depicting the actress saying, "Come over here by me!" was accompanied by the soundtrack, "Where's the baby

going?” and vice versa. This presentation provided a somewhat stringent test of redundancy/synchrony detection in that the audiovisual onset and offset synchrony were preserved in both conditions (the soundtracks to both occurred only while the faces were visible and moving rather than beginning before or terminating after the movement in the asynchronous condition) and only the internal temporal synchrony of the movements of speech with respect to the temporal structure of the sounds of speech was incongruent during asynchronous presentations. All stimuli were 1700 ms in duration and subtended a 33° vertical by 39° horizontal visual angle. The audiovisual stimuli were 60 dB at the position of the infant during testing. The videos consisted of close-up footage of the actress’ face (from the neck-line up). A single actress, positioned in front of a blue-gray background, was used for all stimuli. The stimuli were drawn from the positive affect subset of stimuli used in Flom and Bahrick (2007).

**Sesame Street characters:** Videos of Sesame Street characters were used as attractor stimuli. The Sesame Street videos covered a 15° square area centered on the monitor.

**Procedure**—Infants were held on a parent’s lap approximately 55 cm from the center of the computer monitor. They were fitted with an EGI sensor net and impedances were measured. The test phase consisted of repeated presentations of the unimodal visual, synchronous audiovisual, and asynchronous audiovisual stimuli. The stimuli were presented for 1700 ms, followed by a blank blue-gray screen with a random duration of 950 to 1200 ms. Stimulus type presentations were equally distributed across trials in random order. Stimulus presentations were initiated only when the infant was judged to be looking at the monitor. During periods of distraction, the Sesame Street videos were presented as an attractor stimulus, subsequent stimulus presentations were always preceded by a blank screen for at least 500 ms. The procedure was continued for as long as the infant was not tired or fussy (approximately 10 min on average).

**Fixation Judging**—In addition to judging infant fixations online for the purpose of experimental control during testing, fixations were also judged offline by a trained rater to determine if the participant was looking during each ERP trial. ERP trials in which the infant was not looking at any point during the stimulus presentation were not included in analyses.

**EEG recording and analyses**—The Electrical Geodesics Incorporated (EGI) Geodesic EEG System 300 (GES 300) 128 channel EEG recording system was used. The EGI Netstation program was used for A/D sampling, data storage, zero and gain calibration for each channel, and measuring impedances. Electrodes were adjusted until impedance values ranging from 10 to 50 k $\Omega$  were achieved. The Netstation program received serial communication from a Dell Workstation used to control the experimental protocol with E-Prime 2.0 software (Psychology Software Tools, Inc.). The sampling rate of the EEG was 250 Hz (4 ms samples) and band-pass filters were set from 0.1 to 100 Hz, with 20K amplification. EEG recordings were referenced to the vertex and algebraically re-referenced to the average reference.

The EEG recordings were inspected for artifacts (i.e., blinks, saccades, movement artifact, and drift) and poor recordings using the Netstation review system. Individual channels were marked bad within trials if these occurred. Segments in which more than 10% of the channels were marked bad were eliminated from the analysis. For trials that were retained for the ERP analysis, individual channels marked bad were replaced using a spherical spline interpolation (Perrin, Pernier, Bertrand, Giard, & Echallier, 1987; Srinivasan, Tucker, & Murias, 1998). Only those participants who retained enough ERP trials per condition (i.e., 10 trials) for stable ERP averages following EEG editing were included in the final dataset



(DeBoer, Scott, & Nelson, 2007). The number of trials included in the averages did not differ significantly ( $p > .10$ ) across stimulus types ( $M_s = 16.4$  asynchronous audiovisual, 16.9 synchronous audiovisual, and 14.9 unimodal visual).

ERP averages were calculated from 200 ms before stimulus onset through 1.75 s after stimulus onset. For increased stability, we analyzed the ERP averaged across multiple channels. Nc peak amplitude and latency to peak were analyzed from 350 – 750 ms following stimulus onset at midline frontal (4, 10, 11, 16, 18, 19), midline central (7, 31, 55, 80, 106), and midline parietal (61, 62, 67, 72, 77, 78) electrode locations. For the N290 component, mean amplitude from 190 – 290 ms following stimulus onset was analyzed at left occipital (65, 69, 70), midline occipital (74, 75, 82), and right occipital (83, 89, 90) electrode clusters. For the P400 component, we analyzed mean amplitude from 300 – 500 ms following stimulus onset examining the same electrode locations as the N290 analysis. Electrodes were chosen for the analyses based on past research in the area and visual inspection of the grand average ERP waveforms (DeBoer, Scott, & Nelson, 2007).

**Design for Statistical Analysis**—The design included the experimental factors of stimulus type (unimodal visual, synchronous bimodal, asynchronous bimodal) and electrode location (level varied by component) as repeated measures. Repeated-measures ANOVAs were used in all analyses and the Greenhouse-Geisser correction was used in cases of violations of the assumption of sphericity. For significant effects, follow-up analyses were done using one-way ANOVAs and paired-samples  $t$ -tests. Effect sizes ( $\eta_p^2$ ) are reported on all significant effects, and all significant tests are reported at  $p < .05$ .

## Results

**Primary Analyses: The Nc Component**—Our primary analyses assessed the salience of intersensory redundancy as reflected by the Nc component. The amplitude of Nc component was expected to be higher for redundant (synchronous) than nonredundant (both asynchronous and unimodal visual) stimuli if intersensory redundancy is the basis for the salience of multimodal stimulation in early development. To analyze peak (minimum) amplitude of Nc, we conducted a two-way ANOVA with electrode location (3: midline frontal, midline central, midline parietal) and stimulus type (3: unimodal visual, synchronous bimodal, asynchronous bimodal) as within-subjects factors. There was a significant main effect for electrode location,  $F(2, 28) 22.90$ ;  $p < .001$ ,  $\eta_p^2 = .621$ , with greater amplitude Nc at parietal electrodes than central and frontal electrodes. This main effect was qualified by a significant electrode by stimulus type interaction,  $F(4, 56) 5.12$ ;  $p < .001$ ,  $\eta_p^2 = .268$ . A follow-up ANOVA on parietal electrodes revealed a significant main effect for stimulus type. Consistent with our predictions, infants demonstrated greater amplitude Nc to synchronous audiovisual ( $M = -16.44 \mu V$ ) than asynchronous audiovisual ( $M = -13.10 \mu V$ ,  $p = .019$ ) and unimodal visual ( $M = -11.99 \mu V$ ,  $p = .028$ ) stimuli (see Figure 1).

We analyzed latency to peak for the Nc component using the same statistical approach as above and found similar effects. There was a significant interaction of electrode and stimulus type,  $F(4, 56) 3.34$ ;  $p = .016$ ,  $\eta_p^2 = .193$ . At parietal electrodes, infants demonstrated shorter latency to peak Nc for multimodal stimulus presentations ( $M = 490.76$  ms and 489.50 ms for asynchronous and synchronous respectively) compared to unimodal visual presentations ( $M = 558.04$  ms;  $p < .05$  for both comparisons).

**Secondary Analyses**—We conducted secondary analyses of ERP components involved in face and speech processing. After visual inspection of the grand average waveforms, we focused these analyses on the N290 and P400 components involved in face processing in

infancy, and the auditory P1 involved in speech processing. The N290 and P400 were analyzed at occipital electrodes, and the auditory P1 was analyzed at anterior temporal electrode sites. We predicted significant differences between both audiovisual conditions compared to the unimodal visual condition due to the additive effects of combining auditory and visual stimuli.

**The N290 Component:** We analyzed the average amplitude and peak latency of the N290 using a two-way ANOVA with electrode location (3: left occipital, midline occipital, right occipital) and stimulus type (3: unimodal visual, synchronous audiovisual, asynchronous audiovisual) as within-subjects factors. For mean amplitude, there were significant main effects of electrode location,  $F(1, 28) 5.92; p < .01, \eta_p^2 = .297$ , and stimulus type,  $F(2, 28) 5.02; p = .016, \eta_p^2 = .264$ . Infants demonstrated greater amplitude N290 at central occipital electrodes ( $M = -9.19 \mu V$ ) than left occipital ( $M = -3.46 \mu V; p < .01$ ) and right occipital ( $M = -4.96 \mu V; p < .05$ ) electrodes. Infants also demonstrated greater amplitude N290 on multimodal trials than unimodal visual trials ( $p < .05$  for both comparisons; see Figure 2). No differences were found between the synchronous and asynchronous audiovisual conditions, and there were no significant effects related to stimulus type for latency to peak for the N290 component.

**The P400 Component:** We analyzed the mean amplitude and latency to peak of the P400 using the same statistical approach as our analysis of the N290 component. For amplitude analyses, there was a significant main effect of stimulus type,  $F(2, 28) 5.94; p < .01, \eta_p^2 = .298$ . Similar to the N290 effect, infants demonstrated greater negativity in P400 amplitude on multimodal trials when compared to unimodal visual trials ( $p < .05$  for both comparisons; see figure 2). However, inspection of Figure 2 indicates that these differences were possibly due to the N290 effect as the amount of change occurring in the waveforms from the peak of the N290 to the peak of the P400 was similar across conditions. Thus, we conducted a follow-up analysis examining the peak-to-peak change in amplitude that occurred from the peak of the N290 to the peak of the P400. In the peak-to-peak analysis, no differences were found based on stimulus type. There were no significant effects for latency to peak for the P400 component.

**The Auditory P1 component:** We analyzed mean amplitude and latency to peak of the Auditory P1 from 190 – 390 ms at left anterior temporal (34, 35, 39, 40, 41), and right anterior temporal (103, 109, 110, 115, 116) electrodes. There was a main effect for stimulus type,  $F(2, 28) 5.78; p < .01, \eta_p^2 = .292$ . Infants demonstrated greater amplitude auditory P1 for both asynchronous audiovisual ( $M = 6.00 \mu V, p = .005$ ) and synchronous audiovisual ( $M = 5.45 \mu V, p = .038$ ) than to unimodal visual stimuli ( $M = 1.00 \mu V$ ). No differences were found between the asynchronous and synchronous audiovisual conditions (see Figure 3).

## Discussion

Infants were exposed to redundant (synchronous) audiovisual, nonredundant (asynchronous) audiovisual, and unimodal visual presentations of a woman speaking, and ERP components associated with attention, face processing, and auditory processing were examined. Our main hypothesis for Experiment 1, consistent with the IRH, was that the salience of multimodal stimulation was based on intersensory redundancy and this attentional salience would be reflected by the Nc component. Thus, we predicted that the Nc component would be greater in amplitude following synchronous audiovisual presentations when compared to asynchronous audiovisual and unimodal visual presentations. This prediction was supported by an interaction of electrode and stimulus type on Nc amplitude. At midline parietal electrodes, infants demonstrated greater amplitude Nc in the synchronous audiovisual condition than the asynchronous audiovisual and unimodal visual conditions. Additionally,

the latency to peak of the Nc component was shorter for both audiovisual conditions than the unimodal visual condition. These findings indicate greater sensitivity to multimodal presentations than unimodal presentations, and greater allocation of attention to redundant multimodal than to nonredundant multimodal and unimodal stimuli. These findings provide novel information about neural mechanisms underlying the facilitating effects of intersensory redundancy on infant attention.

Descriptive analyses of face processing components revealed that infants demonstrated greater amplitude of the N290 component following synchronous and asynchronous audiovisual presentations compared to unimodal visual presentations. While the results of our N290 analysis indicate that multimodal stimulation (regardless of synchrony) may enhance face processing in infants, we cannot rule out the possibility that this multimodal effect is simply due to linear super-position of the electrical activity associated with auditory and visual stimulation as opposed to enhanced responsiveness. Without a unimodal auditory condition, we cannot determine if this effect is super-additive. We chose not to include a unimodal auditory condition because it is generally advised that researchers limit the number of stimulus types to two or three in infant ERP research to avoid excessively high attrition rates (e.g., DeBoer et al., 2007). Infants also demonstrated greater amplitude auditory P1 in both audiovisual conditions compared to the unimodal visual condition. This was expected given the lack of auditory stimulation in the unimodal visual condition.

## **Experiment 2: Intersensory Redundancy and Processing Efficiency as assessed by the LSW and Nc Component**

The findings from Experiment 1 indicate that redundant audiovisual stimuli elicit greater amplitude Nc than nonredundant audiovisual and unimodal visual stimuli. This enhanced neural responsiveness associated with attention may serve as a neural mechanism underlying the intersensory facilitation of perceptual learning and recognition memory that has been consistently found in infant habituation studies (e.g., Bahrick, Flom, & Lickliter, 2002; Bahrick & Lickliter, 2000; Flom & Bahrick, 2007). Experiment 2 was designed to examine the influence of redundant and nonredundant audiovisual stimuli on the LSW associated with stimulus processing (i.e., encoding) and recognition memory in infancy. A reduction in the amplitude of the LSW with repeated stimulus exposure is associated with recognition memory of a fully processed stimulus (de Haan, 2007; de Haan & Nelson, 1997, 1999; Reynolds, Guy, & Zhang, 2011). For example, Snyder (2010) found that infants who demonstrate a significant reduction in LSW amplitude at anterior temporal electrodes following repeated presentation of a single stimulus were more likely to show evidence of recognition memory for the previously viewed stimulus in behavioral testing than infants who showed no reduction in LSW amplitude.

In Experiment 2, we exposed 5-month-old infants to repeated presentations of a single stimulus (either synchronous audiovisual or asynchronous audiovisual) and utilized a block design (where the same trial type was presented across three blocks) to allow for comparison of the amplitude of the LSW across early to late trials. We utilized a between-subjects design and only presented a single stimulus to infants in each group to avoid potential interference effects from other stimulus types. Consistent with behavioral findings indicating enhanced perceptual processing and learning of redundant multimodal stimuli (e.g., Bahrick & Lickliter, 2000, 2002, 2012), we predicted that infants would demonstrate reduced amplitude LSWs across early to late trials in the synchronous audiovisual condition, but that no differences would be found in LSW amplitude across early to late trials in the asynchronous audiovisual condition. Based on the results of Experiment 1, we also predicted that infants would demonstrate greater amplitude Nc in the synchronous audiovisual condition than the asynchronous audiovisual condition. Taken together, these findings

would reveal neural underpinnings of greater attention to and enhanced processing of redundant audiovisual stimuli compared to nonredundant audiovisual stimuli.

## Method

**Participants**—Twenty-two 5-month-old infants (10 male, 12 female) were tested. Recruitment and inclusion criteria were the same as for Experiment 1. The ethnic/racial distribution of participants was: 20 Caucasian (not Hispanic), 1 Hispanic, and 1 Biracial. An additional 20 infants were tested, but not included in the final sample due to fussiness/distractibility ( $N = 9$ ), excessive artifact in the EEG ( $N = 9$ ), and technical problems ( $N = 2$ ).

**Stimuli**—The stimuli were identical to those used in Experiment 1 with the exception that infants were only exposed to a single test stimulus as opposed to 6 test stimuli (i.e., 2 exemplars from 3 stimulus types in Experiment 1). Infants were exposed to either an exemplar from the synchronous audiovisual stimulus type or an exemplar from the asynchronous audiovisual stimulus type.

**Procedure**—The procedure was identical to Experiment 1 with the exception that participants were shown only a single, repeated stimulus, with half the participants receiving a synchronous audiovisual stimulus and half an asynchronous audiovisual stimulus. A block design was utilized. Infants were exposed to a total of 90 stimulus presentations consisting of 3 blocks of 30 trials each. The procedure lasted approximately 10 minutes.

**Fixation Judging**—Fixation Judging was done in the same manner as Experiment 1.

**EEG recording and analyses**—The general approach to EEG recording was the same as that used in Experiment 1. Only those participants who retained enough ERP trials per block (i.e., 10 trials) for stable ERP averages following EEG editing were included in the final dataset. Infants were more likely to become bored or fussy toward the end of the procedure, thus only a few infants contributed 10 artifact-free trials on the third block (trials 60 through 90). To avoid unreasonably high attrition, blocks 2 and 3 were combined into a “late” block for comparison with the “early” block of trials (i.e., block 1 – trials 1 through 30). Because the number of trials included in ERP averages can affect the amplitude of the ERP waveform, equal numbers of trials were included in the averages for the early and late blocks per each participant. For example, if an infant contributed 20 good trials to the ERP average for the early block (block 1), the first 20 artifact-free trials from blocks 2 and 3 were used in that participant’s ERP average for the late block. Using this blocking procedure, the average number of trials per block was 19.64 ( $SD = 3.76$ ) for the synchronous audiovisual condition and 20.00 ( $SD = 4.84$ ) for the asynchronous audiovisual condition.

Nc amplitude was examined using the same electrodes and approach used in Experiment 1. For the LSW, mean amplitude from 1 – 1.5 s following stimulus onset was analyzed at left anterior temporal (41, 45, 46, 50, 57), and right anterior temporal (100, 101, 102, 103, 108) electrodes. Electrode locations were chosen for the analysis based on past research and visual inspection of the grand average waveforms.

**Design for Statistical Analysis**—The design for the LSW analysis included stimulus type (synchronous audiovisual, asynchronous audiovisual) as a between-subjects factor, and block (early, late) and electrode location (left temporal, right temporal) as within-subjects factors. A mixed ANOVA was used for a full factorial analysis and follow-up analyses were done using one-way ANOVAs and paired-samples  $t$ -tests.

## Results

**Primary Analyses: The Late Slow Wave**—A three-way ANOVA was conducted on mean amplitude of the LSW with electrode location (left temporal, right temporal) and block (early, late) as within-subjects factors, and stimulus type (synchronous audiovisual, asynchronous audiovisual) as a between-subjects factor. There was a significant interaction between block and electrode location,  $F(1, 20) 5.44$ ;  $p < .03$ ,  $\eta_p^2 = .214$ . Infants demonstrated greater amplitude LSWs,  $t(21) -2.77$ ,  $p = .012$ , at right temporal electrodes during the early block of trials ( $M = -5.08 \mu V$ ) than during the late block of trials ( $M = .45 \mu V$ ). Additionally, in the early block of trials, infants demonstrated greater amplitude LSWs at right temporal ( $M = -5.08 \mu V$ ) compared to left temporal ( $M = .16 \mu V$ ) electrodes,  $t(21) -2.33$ ,  $p = .03$ .

In support of our primary hypothesis that only infants in the synchronous audiovisual condition would demonstrate a reduction in amplitude of the LSW across early to late blocks of trials, there was a significant interaction between block and condition,  $F(1, 20) 4.94$ ;  $p = .04$ ,  $\eta_p^2 = .198$ . Infants in the synchronous audiovisual condition demonstrated a significant reduction,  $t(10) -2.61$ ,  $p = .026$ , in the amplitude of the LSW at temporal electrodes from early ( $M = -4.70 \mu V$ ) to late trials ( $M = -2.65 \mu V$ ). For the asynchronous condition, no differences,  $t(10) = .296$ ;  $p = .773$ , were found in the amplitude of the LSW across early ( $M = -.22 \mu V$ ) to late trials ( $M = -.62 \mu V$ ). Furthermore, in the early block of trials, the amplitude of the LSW was significantly higher in the synchronous condition than in the asynchronous condition,  $t(20) = 2.73$ ,  $p = .01$ . No differences in LSW amplitude were found across conditions in the late block of trials. Interestingly, when analyzing the LSW separately within each hemisphere, infants in the synchronous audiovisual condition (see top right panel of Figure 4) showed a significant reduction in LSW amplitude from early to late trials at right temporal electrodes,  $t(10) -2.75$ ,  $p = .02$ . In contrast, infants in the asynchronous audiovisual condition (see bottom left panel of Figure 4) showed a non-significant increase in LSW amplitude from early to late trials at left temporal electrodes,  $t(10) 1.993$ ,  $p = .07$ . No differences were found across early to late trials for any other comparisons (all  $ps > .10$ ).

**Secondary Analyses: The Nc Component**—A three-way ANOVA was done on Nc peak amplitude with electrode location (frontal, central, parietal) and block (early, late) as within-subjects factors, and stimulus type (synchronous audiovisual, asynchronous audiovisual) as a between-subjects factor. There was a significant main effect for electrode location,  $F(2, 40) 49.20$ ;  $p < .001$ ,  $\eta_p^2 = .711$ , similar to Experiment 1. Nc was greater in amplitude at midline parietal electrodes than midline central and midline frontal electrodes. Replicating the results of Experiment 1, planned comparisons revealed a significant effect of stimulus type on Nc amplitude at midline frontal and central electrodes,  $F(1, 20) 4.63$ ;  $p = .044$ ,  $\eta_p^2 = .188$ , with infants demonstrating greater amplitude Nc to synchronous audiovisual ( $M = -6.79 \mu V$ ) compared to asynchronous audiovisual stimuli ( $M = -1.74 \mu V$ ),

## Discussion

Experiment 2 examined the influence of redundant (synchronous) and nonredundant (asynchronous) audiovisual stimuli on attention and stimulus processing in 5-month-old infants. We predicted that intersensory redundancy would be associated with a reduction in amplitude of the LSW across repeated stimulus presentations and with greater amplitude Nc. Our findings supported both of these predictions. Infants in the synchronous audiovisual condition showed a reduction in amplitude of the LSW across early to late trials, which is indicative of effective stimulus encoding resulting in recognition of the repeated stimulus (e.g., de Haan, 2007; Snyder, 2010; Reynolds, Guy, & Zhang, 2011). No significant differences were found in LSW amplitude across early to late trials for infants in the

asynchronous audiovisual condition. These findings indicate that infants processed redundant audiovisual stimuli more efficiently than nonredundant audiovisual stimuli. In addition, infants demonstrated greater amplitude Nc in the synchronous condition compared to the asynchronous condition. This finding replicates the Nc results from Experiment 1, and is consistent with past research demonstrating greater amplitude Nc in response to congruent face-voice pairings compared to incongruent face-voice pairings (Vogel et al., 2012). Taken together, these findings indicate that redundant audiovisual stimulation is associated with greater attention allocation (Experiment 1) and more efficient stimulus processing (Experiment 2) than nonredundant audiovisual stimulation.

## General Discussion

The findings from the current study are the first to demonstrate support for the intersensory redundancy hypotheses at the neural level in human infants, and reveal new information above and beyond that provided by the behavioral literature about the mechanism underlying enhanced processing of intersensory redundancy in early infancy. Attentional salience was indexed using amplitude of the Nc component. Infants demonstrated greater amplitude of the Nc component to synchronous audiovisual stimulation compared to asynchronous audiovisual stimulation in both Experiments 1 and 2. These stimulus types provided the same amount and type of stimulation and differed only in terms of intersensory redundancy, the temporal relations between the audible and visual stimulation. This finding thus indicates that intersensory redundancy per se is salient to infants. Infants also showed greater Nc amplitude in synchronous audiovisual as compared with unimodal visual stimulation in Experiment 1. The Nc component is ubiquitous in the infant ERP literature and has been found to be associated with visual attention and stimulus salience (e.g., Ackles, 2008; Courchesne et al., 1981; Reynolds & Richards, 2005; Reynolds, Courage, & Richards, 2010; Richards, 2003). For example, infants demonstrate greater Nc amplitude to stimuli they visually prefer and increased amplitude of Nc has been proposed to reflect activation of a general arousal system involved in attention (Reynolds, Courage, & Richards, in press). The finding that Nc is greater in amplitude when heart rate measures are indicative of attention (Reynolds et al., 2010; Richards, 2003) provides further support for this proposal.

Efficiency of stimulus processing was indexed through analyzing changes in the amplitude in the LSW across blocks of trials indicative of recognition memory. The LSW has been proposed to reflect stimulus encoding or an updating of working memory for a partially processed stimulus (de Haan & Nelson, 1997). Thus, a reduction of the LSW with repeated exposure is indicative of successful encoding and recognition of the repeated or familiar stimulus (e.g., Snyder, 2010). Infants in Experiment 2 were exposed to a single stimulus across trial blocks, composed of either synchronous or asynchronous audiovisual speech. If intersensory redundancy promotes enhanced perceptual processing, then infants in the synchronous but not the asynchronous audiovisual condition were expected to show a significant reduction in the amplitude of the LSW across early to late trials. An interaction of stimulus condition by block confirmed this prediction. Only infants in the synchronous condition demonstrated a significant reduction in the amplitude of the LSW across early to late trials. No differences were found in LSW amplitude across early to late trials for infants in the asynchronous audiovisual condition. The current findings indicate that intersensory redundancy available in multimodal stimulation is not only salient but promotes more efficient perceptual processing than nonredundant stimulation. These findings converge with those of behavioral studies using habituation with similar stimuli. For example, Flom & Bahrick (2007) have shown that intersensory redundancy facilitates discrimination of affect in videos of women speaking, with discrimination demonstrated 3 months earlier in development (i.e., at 4 months as opposed to 7 months) in synchronous audiovisual speech as compared with unimodal visual speech.

One other ERP study (Hyde, et al 2011) also found that 5-month-old infants discriminated synchronous and asynchronous audiovisual stimuli. In contrast to our findings, infants in their study showed greater amplitude Nc to asynchronous audiovisual stimuli than to synchronous audiovisual stimuli. The authors concluded that the greater amplitude Nc to asynchronous audiovisual stimuli reflected detection of a novel stimulus category (consistent with early interpretations that the Nc component is associated with novelty detection, e.g., Courchesne et al., 1981). However, due to the lack of a balanced design in their stimulus set (i.e. the asynchronous condition provided a visual stimulus not used in the synchronous condition), their findings may have been driven by oddball effects on Nc amplitude as opposed to asynchrony per se. Hyde and colleagues (2011) also did not report the number of trials included in their ERP averages across experimental conditions, a factor which is known to affect peak amplitude and the signal-to-noise ratio of the averaged waveform (Luck, 2005). Thus, due to a number of potential confounds, the results of their analysis of the effects of intersensory redundancy on infant attention remain inconclusive. In a recent study (Reynolds, Zhang, & Guy, in press) examining look duration in 3-, 6-, and 9-month-old participants, infants looked significantly longer at both synchronous and asynchronous audiovisual stimuli in comparison to unimodal visual stimuli. No differences were found across audiovisual conditions. Taken together, these findings indicate that both types of multimodal stimuli are highly salient in infancy.

In the current study, we utilized a fully balanced stimulus set to control for potential oddball (or frequency) effects. Asynchrony was achieved, similar to Hyde et al, (2011), by playing the soundtrack to one phrase while presenting the visual speech of another phrase. However in our study, the audio and visual phrases were counterbalanced so that infants received both phrases in both the synchronous and asynchronous conditions. Infants demonstrated greater amplitude Nc to synchronous audiovisual stimuli compared to asynchronous audiovisual stimuli in both Experiments 1 and 2. These findings are consistent with behavioral evidence of the attentional salience of intersensory redundancy as well as recent fMRI work with adults (Marchant, Ruff, & Driver, 2012) demonstrating significantly greater BOLD response to synchronous audiovisual stimuli compared to asynchronous audiovisual stimuli.

The use of a block design in Experiment 2 allowed us to directly examine the impact of synchronous and asynchronous audiovisual stimuli on the efficiency of stimulus processing by assessing change in the LSW across blocks. There was a significant interaction between stimulus condition and block. Only infants in the synchronous condition showed a significant reduction in the amplitude of the LSW over time. No significant differences in LSW amplitude were found across early to late trials in the asynchronous condition. There was also an interaction between electrode location and block. In the early block of trials, the LSW was greater in amplitude at right temporal electrodes than at left temporal electrodes. Additionally, the reduction in amplitude of the LSW from early to late blocks for the synchronous groups was most evident at right temporal electrodes (see figure 4). These findings are consistent with past work indicating the LSW to faces is often lateralized over right hemisphere electrode sites (de Haan, Johnson, & Halit, 2007). Taken together, these results provide evidence of intersensory facilitation of attention to and processing of face-voice pairings for 5-month-old infants occurring with exposure to redundant audiovisual stimuli.

In sum, findings from the present study provide new information about the neural processes underlying the enhanced processing of redundant multimodal stimuli in infancy. The enhanced responsiveness of infants to redundant audiovisual stimuli begins at the level of attention, as indicated by greater amplitude Nc to synchronous audiovisual stimuli compared to asynchronous audiovisual and unimodal visual stimuli. This enhanced attention response is followed by changes over time in the LSW that are indicative of recognition memory for a

fully processed stimulus. The current findings converge with those of behavioral studies to demonstrate that naturalistic, synchronous, multimodal events are highly salient because they provide intersensory redundancy, which attracts infant attention to amodal stimulus properties and enhances information processing during early development.

## Acknowledgments

Notes: The authors wish to thank Dantong Zhang for her assistance with data collection and data processing. We are especially grateful to the parents and infants who participated in this study. Support for this research was provided in part by NICHD grant R03 HD05600 awarded to GR; NICHD grants K02 HD064943 and RO1 HD053776 awarded to LB, and NSF grant BCS1057898 awarded to RL.

## References

- Ackles PK. Stimulus novelty and cognitive-related ERP components of the infant brain. *Perceptual and Motor Skills*. 2008; 106:3–20. [PubMed: 18459351]
- Bahrick LE, Flom R, Lickliter R. Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Developmental Psychobiology*. 2002; 41:352–363. [PubMed: 12430159]
- Bahrick LE, Lickliter R. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*. 2000; 36(2):190–201. [PubMed: 10749076]
- Bahrick, LE.; Lickliter, R. Intersensory redundancy guides early perceptual and cognitive development. In: Kail, R., editor. *Advances in child development and behavior*. Vol. 30. New York: Academic Press; 2002. p. 153-187.
- Bahrick, LE.; Lickliter, R. The role of intersensory redundancy in early perceptual, cognitive, and social development. In: Bremner, A.; Lewkowicz, DJ.; Spence, C., editors. *Multisensory development*. Oxford, England: Oxford University Press; 2012. p. 183-206.
- Bahrick, LE.; Pickens, J. Amodal relations: The basis for intermodal perception and learning in infancy. In: Lewkowicz, DJ.; Lickliter, R., editors. *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum; 1994. p. 205-233.
- Beauchamp MS. Statistical criteria in fMRI studies of multisensory integration. *Neuroinformatics*. 2005; 3:93–113. [PubMed: 15988040]
- Benasich AA, Choudhury N, Friedman JT, Realpe-Bonilla T, Chojnowska C, Gou Z. The infant as a prelinguistic model for language learning impairments: predicting from event-related potentials to behavior. *Neuropsychologia*. 2006; 44:396–411. [PubMed: 16054661]
- Benevento LA, Fallon J, Davis BJ, Rezak M. Auditory-visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. *Experimental Neurology*. 1977; 57:849–872. [PubMed: 411682]
- Bremner, A.; Lewkowicz, DJ.; Spence, C. *Multisensory development*. New York: Oxford University Press; 2012.
- Bristow D, Dehaene-Lambertz G, Mattout J, Soares C, Gliga T, Baillet S, Magin J. Hearing faces: How the infant brain matches the face it sees with the speech it hears. *Journal of Cognitive Neuroscience*. 2008; 21:905–921. [PubMed: 18702595]
- Bruce C, Desimone R, Gross CG. Visual properties of neurons in a polysensory area in the superior temporal sulcus of the macaque. *Journal of Neurophysiology*. 1981; 46(2):369–384. [PubMed: 6267219]
- Calvert, G.; Spence, C.; Stein, BE. *Handbook of multisensory processes*. Cambridge, MA: MIT Press; 2004.
- Cohen LB. Attention-getting and attention-holding processes of infant visual preferences. *Child Development*. 1972; 43:869–879. [PubMed: 5056611]
- Courchesne E. Event-related brain potentials: Comparison between children and adults. *Science*. 1977; 197:589–592. [PubMed: 877575]
- Courchesne E, Ganz L, Norcia AM. Event-related brain potentials to human faces in infants. *Child Development*. 1981; 52:804–811. [PubMed: 7285651]



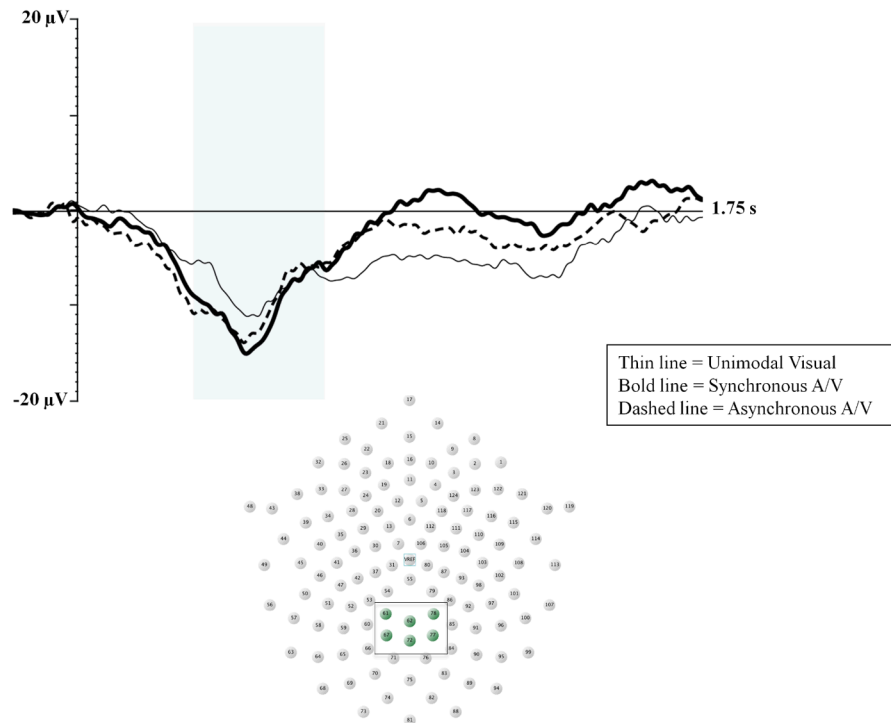
- de Haan, M. Visual attention and recognition memory in infancy. In: de Haan, M., editor. *Infant EEG and event-related potentials*. New York: Psychology Press; 2007. p. 101-144.
- de Haan, M.; Johnson, MH.; Halit, H. Development of face-sensitive event-related potentials during infancy. In: de Haan, M., editor. *Infant EEG and event-related potentials*. New York: Psychology Press; 2007. p. 77-99.
- de Haan M, Nelson CA. Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Child Development*. 1997; 68:187-210. [PubMed: 9179998]
- de Haan M, Nelson CA. Brain activity differentiates face and object processing in 6-month-old infants. *Developmental Psychology*. 1999; 35:1113-1121. [PubMed: 10442879]
- DeBoer, T.; Scott, LS.; Nelson, CA. Methods for acquiring and analyzing infant event-related potentials. In: de Haan, M., editor. *Infant EEG and event-related potentials*. New York: Psychology Press; 2007. p. 5-37.
- deRegnier RA, Nelson CA, Thomas KM, Wewerka S, Georgieff MK. Neurophysiologic evaluation of auditory recognition memory in healthy newborn infants and infants of diabetic mothers. *Journal of Pediatrics*. 2000; 137:777-784. [PubMed: 11113833]
- deRegnier RA, Wewerka S, Georgieff MK, Mattia F, Nelson CA. Influences of post-conceptual age and postnatal experience on the development of auditory recognition memory in the newborn infant. *Developmental Psychobiology*. 2002; 41:216-225. [PubMed: 12325136]
- Farroni T, Csibra G, Simion F, Johnson MH. Eye contact detection in humans from birth. *PNAS*. 2002; 99(14):9602-9605. [PubMed: 12082186]
- Flom R, Bahrick LE. The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology*. 2007; 43:238-252. [PubMed: 17201522]
- Flom R, Bahrick LE. The effects of intersensory redundancy on attention and memory: Infants' long-term memory for orientation in audiovisual events. *Developmental Psychology*. 2010; 46(2):428-436. [PubMed: 20210501]
- Ghazanfar AA, Schroeder CE. Is neocortex essentially multisensory? *Trends in Cognitive Science*. 2006; 10:278-285.
- Giard MH, Peronnet F. Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*. 1999; 11:473-490. [PubMed: 10511637]
- Gibson, EJ.; Pick, AD. *An ecological approach to perceptual learning and development*. New York: Oxford University Press; 2000.
- Gobbelé R, Schürmann M, Forss N, Juottonen K, Buchner H, Hari R. Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *NeuroImage*. 2003; 20:503-511. [PubMed: 14527610]
- Grossmann T, Striano T, Friederici AD. Crossmodal integration of emotional information from face and voice in the infant brain. *Developmental Science*. 2006; 9:309-315. [PubMed: 16669802]
- Halit H, Csibra G, Volein A, Johnson MH. Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*. 2004; 45(7):1228-1234. [PubMed: 15335343]
- Halit H, de Haan M, Johnson MH. Cortical specialization for face processing: Face-sensitive event-related potential components in 3- and 12-month-old infants. *NeuroImage*. 2003; 19:1180-1193. [PubMed: 12880843]
- Hikosaka K. The polysensory region in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Biomedical Research (Tokyo)*. 1993; 14:41-45.
- Hyde DC, Jones BL, Flom R, Porter CL. Neural signatures of face-voice synchrony in 5-month-old human infants. *Developmental Psychobiology*. 2011; 53(4):359-370. [PubMed: 21271561]
- Hyde DC, Jones BL, Porter CL, Flom R. Visual stimulation enhances auditory processing in 3-month-old infants and adults. *Developmental Psychobiology*. 2010; 52(2):181-189. [PubMed: 20014224]
- Jay MF, Sparks DL. Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature*. 1984; 309:345-347. [PubMed: 6727988]
- Karrer, R.; Ackles, PK. Visual event-related potentials of infants during a modified oddball procedure. In: Johnson, R.; Rohrbaugh, JW.; Parasuraman, R., editors. *Current trends in event-related potential research*. Amsterdam: Elsevier Science Publishers; 1987. p. 603-608.

- Karrer, R.; Ackles, PK. Brain organization and perceptual/cognitive development in normal and Down syndrome infants: A research program. In: Vietze, P.; Vaughan, HG., Jr, editors. *The early identification of infants with developmental disabilities*. Philadelphia: Grune & Stratton; 1988. p. 210-234.
- Karrer R, Monti LA. Event-related potentials of 4–7 week-old infants in a visual recognition memory task. *Electroencephalography and Clinical Neurophysiology*. 1995; 94:414–424. [PubMed: 7607095]
- Kushnerenko E, Teinonen T, Volein A, Csibra G. Electrophysiological evidence of illusory audiovisual speech percept in human infants. *Proceedings of the National Academy of Sciences of the USA*. 2008; 105:11442–11445. [PubMed: 18682564]
- Laurienti PJ, Wallace MT, Maldjian JA, Susi CM, Stein BE, Burdette JH. Cross-modal sensory processing in the anterior cingulate and medial prefrontal cortices. *Human Brain Mapping*. 2003; 19:213–233. [PubMed: 12874776]
- Lewkowicz DJ. The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*. 2000; 126:281–308. [PubMed: 10748644]
- Lewkowicz, DJ.; Lickliter, R. *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum; 2004.
- Lickliter R, Bahrick LE. The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin*. 2000; 126:260–280. [PubMed: 10748643]
- Linden JF, Grunewald A, Andersen RA. Responses to auditory stimuli in macaque lateral intraparietal area II. behavioral modulation. *Journal of Neurophysiology*. 1999; 82:343–358. [PubMed: 10400963]
- Luck, SJ. *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press; 2005.
- Lütkenhöner B, Lammertmann C, Simoes C, Hari R. Magnetoencephalographic correlates of audiotactile interaction. *NeuroImage*. 2002; 15:509–522. [PubMed: 11848694]
- Marchant JL, Ruff CC, Driver J. Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Human Brain Mapping*. 2012; 33:1212–1224. [PubMed: 21953980]
- Mazzoni P, Bracewell RM, Barash S, Andersen RA. Spatially-tuned auditory responses in area LIP of macaques performing delayed memory saccades to acoustic targets. *Journal of Neurophysiology*. 1996; 75:1233–1241.
- Molholm S, Ritter W, Murray MM, Javitt DC, Schroeder CE, Foxe JJ. Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*. 2002; 14:115–128. [PubMed: 12063135]
- Nelson, CA. Neural correlates of recognition memory in the first postnatal year of life. In: Dawson, G.; Fischer, K., editors. *Human behavior and the developing brain*. New York, NY: Guilford Press; 1994. p. 269-313.
- Nelson CA, Collins PF. Event-related potential and looking-time analysis of infants' responses to familiar and novel events: Implications for visual recognition memory. *Developmental Psychology*. 1991; 27:50–58.
- Nelson CA, Collins PF. Neural and behavioral correlates of visual recognition memory in 4- and 8-month-old infants. *Brain and Cognition*. 1992; 19:105–121. [PubMed: 1605948]
- Nikkel L, Karrer R. Differential effects of experience on the ERP and behavior of 6-month-old infants: Trends during repeated stimulus presentation. *Developmental Neuropsychology*. 1994; 10:1–11.
- Perrin F, Pernier J, Bertrand O, Giard MH, Echallier JF. Mapping of scalp potentials by surface spline interpolation. *Electroencephalography and Clinical Neurophysiology*. 1987; 66(1):75–81. [PubMed: 2431869]
- Quinn PC, Westerlund A, Nelson CA. Neural markers of categorization in 6-month-old infants. *Psychological Science*. 2006; 17:59–66. [PubMed: 16371145]
- Reynolds GD, Courage ML, Richards JE. Infant attention and visual preferences: Converging evidence from behavior, event-related potentials, and cortical source localization. *Developmental Psychology*. 2010; 46:886–904. [PubMed: 20604609]

- Reynolds, GD.; Courage, ML.; Richards, JE. The development of attention. In: Reisberg, D., editor. *Oxford Handbook of Cognitive Psychology*. Oxford: Oxford University Press; (in press)
- Reynolds GD, Guy MW. Brain–behavior relations in infancy: Integrative approaches to examining infant looking behavior and event-related potentials. *Developmental Neuropsychology*. 2012; 37(3):210–225. [PubMed: 22545659]
- Reynolds GD, Guy MW, Zhang D. Neural correlates of individual differences in infant visual attention and recognition memory. *Infancy*. 2011; 16(4):368–391. [PubMed: 21666833]
- Reynolds GD, Richards JE. Familiarization, attention, and recognition memory in infancy: An ERP and cortical source localization study. *Developmental Psychology*. 2005; 41:598–615. [PubMed: 16060807]
- Reynolds GD, Richards JE. Cortical source localization of infant cognition. *Developmental Neuropsychology*. 2009; 34(3):312–329. [PubMed: 19437206]
- Reynolds GD, Zhang D, Guy MW. Infant attention to dynamic audiovisual stimuli: Look duration from 3 to 9 months of age. *Infancy*. (in press).
- Richards JE. Attention affects the recognition of briefly presented visual stimuli in infants: An ERP study. *Developmental Science*. 2003; 6:312–328. [PubMed: 16718304]
- Rivera-Gaxiola M, Klarman L, Garcia-Sierra A, Kuhl PK. Neural patterns to native and non-native speech contrasts in 11 month-old American infants. *NeuroReport*. 2005; 16:495–498. [PubMed: 15770158]
- Rivera-Gaxiola M, Silva-Pereyra J, Kuhl PK. Brain potentials to native and nonnative speech contrasts in 7 and 11 month old American infants. *Developmental Science*. 2005; 8:162–172. [PubMed: 15720374]
- Santangelo V, Van der Lubbe RHJ, Olivetti-Berlardinelli M, Postma A. Multisensory integration affects ERP components elicited by exogenous cues. *Experimental Brain Research*. 2008; 185:269–277. [PubMed: 17909764]
- Senkowski D, Saint-Amour D, Gruber T, Foxe JJ. Look who’s talking: The deployment of visuo-spatial attention during multisensory speech processing under noisy environmental conditions. *Neuroimage*. 2008; 43(2):379–387. [PubMed: 18678262]
- Snyder K. Neural correlates of encoding predict infants’ memory in the paired-comparison procedure. *Infancy*. 2010; 15:487–516.
- Snyder K, Garza J, Zolot L, Kresse A. Electrophysiological signals of familiarity and recency in the infant brain. *Infancy*. 2010; 15:487–516.
- Snyder K, Webb SJ, Nelson CA. Theoretical and methodological implications of variability in infant brain response during a recognition memory paradigm. *Infant Behavior and Development*. 2002; 25:466–494.
- Srinivasan R, Tucker DM, Murias M. Estimating the spatial Nyquist of the human EEG. *Behavioral Research Methods, Instruments, & Computers*. 1998; 30:8–19.
- Stein, BE.; Meredith, MA.; Wallace, M. Development and neural basis of multisensory integration. In: Lewkowicz, DJ.; Lickliter, R., editors. *The development of intersensory perception: Comparative perspectives*. Hillsdale, NJ: Erlbaum; 1994. p. 81-105.
- Stein BE, Stanford TR. Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*. 2008; 9:255–266.
- Vogel M, Monesson A, Scott LS. Building biases in infancy: The influence of race on face and voice emotion matching. *Developmental Science*. 2012; 15:359–372. [PubMed: 22490176]
- Walker-Andrews AS. Infants’ perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*. 1997; 121:437–456. [PubMed: 9136644]
- Wallace MT, Stein BE. Development of multisensory neurons and multisensory integration in cat superior colliculus. *Journal of Neuroscience*. 1997; 17:2429–2444. [PubMed: 9065504]
- Wallace MT, Wilkinson LK, Stein BE. Representation and integration of multiple sensory inputs in primate superior colliculus. *Journal of Neurophysiology*. 1996; 76:1246–1266. [PubMed: 8871234]
- Webb SJ, Long JD, Nelson CA. A longitudinal investigation of visual event-related potentials in the first year of life. *Developmental Science*. 2005; 8:605–616. [PubMed: 16246251]

Wiebe SA, Cheatham CL, Lukowski AF, Haight JC, Muehleck AJ, Bauer PJ. Infants' ERP responses to novel and familiar stimuli change over time: Implications for novelty detection and memory. *Infancy*. 2006; 9:21–44.

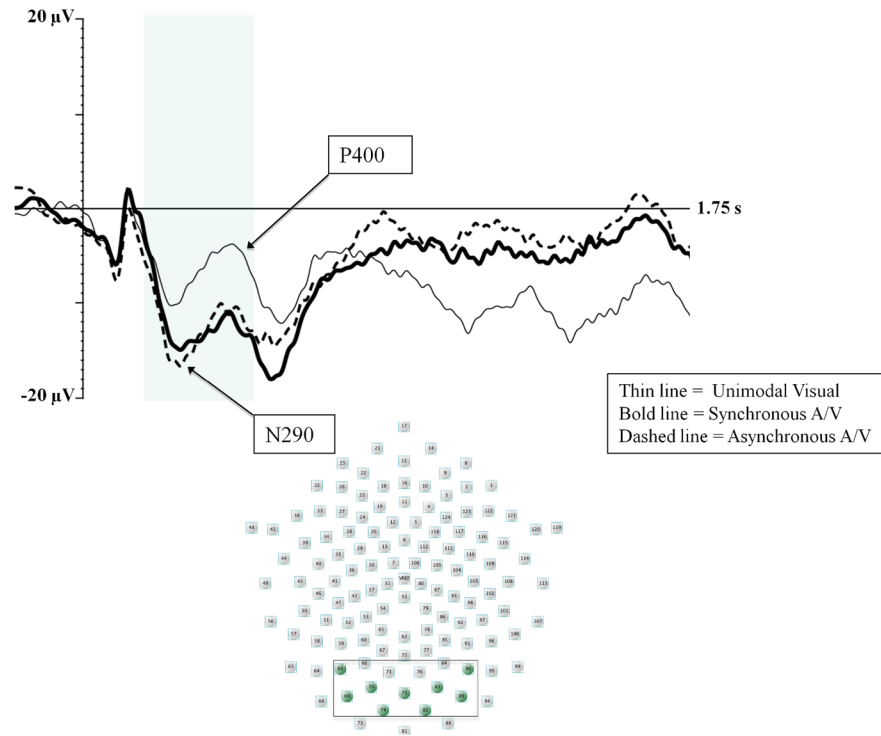
## Nc ERP Component at Midline Parietal Electrodes



**Figure 1.**

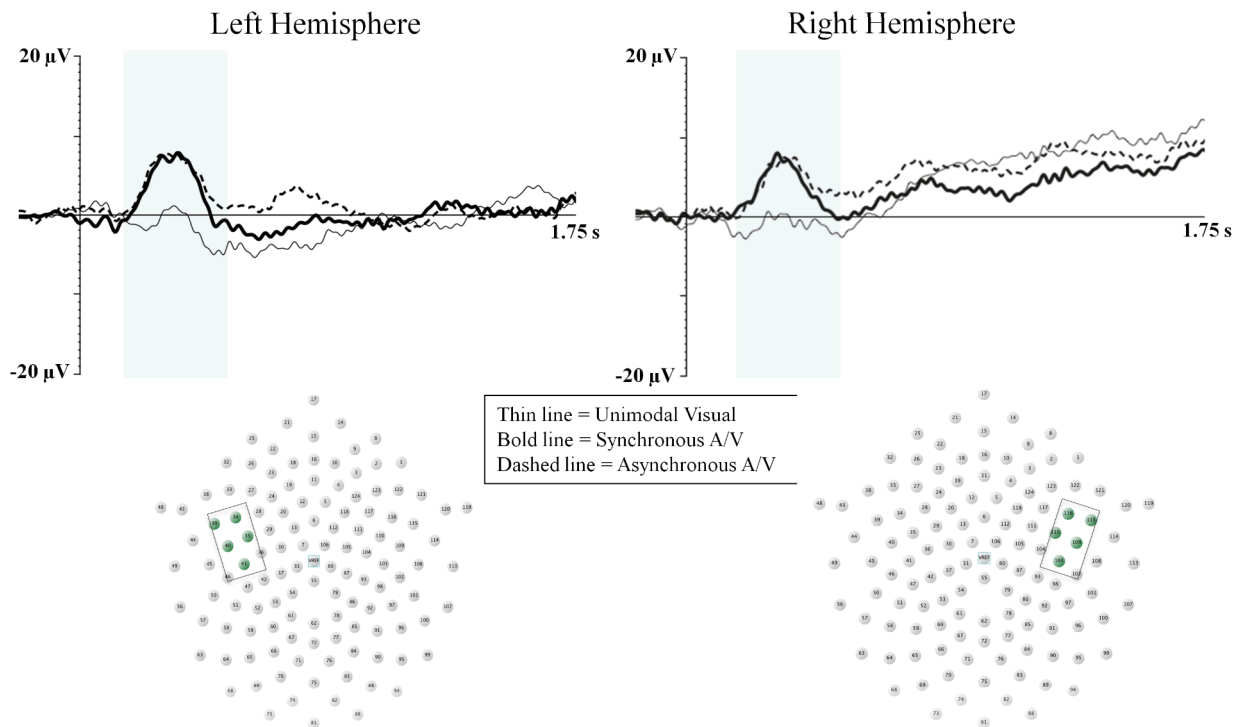
The Nc component at midline parietal electrodes is shown for the unimodal visual (thin line), synchronous audiovisual (bold line), and asynchronous audiovisual (dashed line) conditions from Experiment 1. The Y-axis represents the amplitude of the ERP in microvolts, and the X-axis represents time following stimulus onset. The time-window of the component analysis is shaded on the X-axis. The positioning of the electrodes included in the midline parietal cluster are shown within the electrode montage (see box and shaded electrode sites).

N290 and P400 ERP Components at Occipital Electrodes



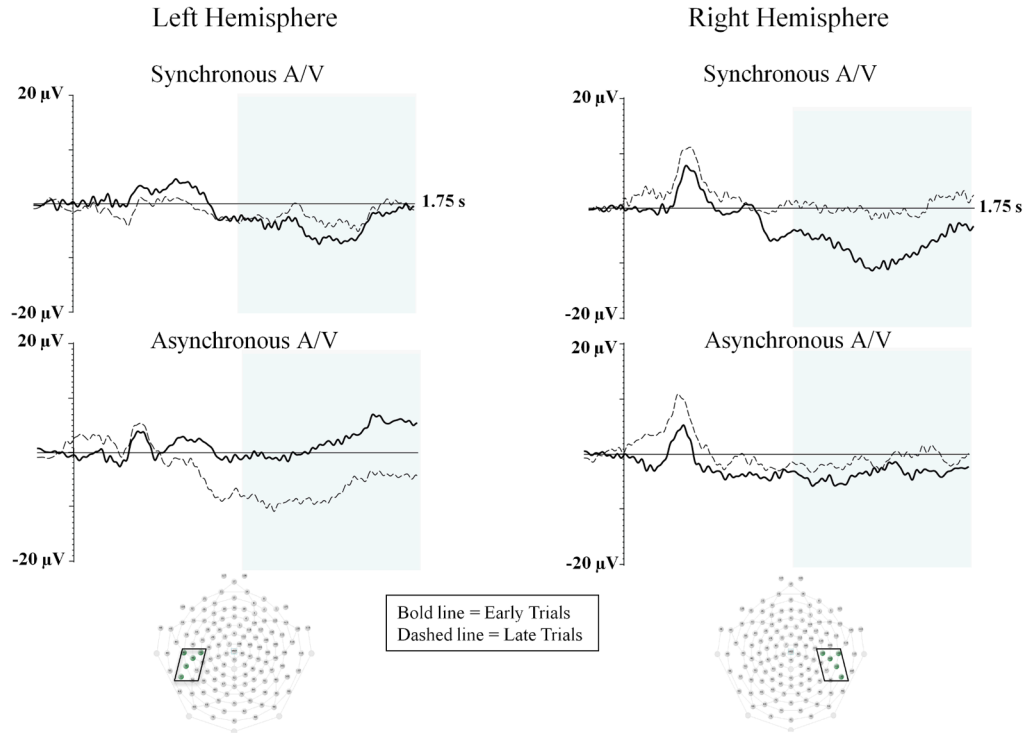
**Figure 2.** The N290 and P400 components are shown for the unimodal visual (thin line), synchronous audiovisual (bold line), and asynchronous audiovisual (dashed line) conditions from Experiment 1. The Y-axis represents the amplitude of the ERP in microvolts, and the X-axis represents time following stimulus onset. The time-window of the component analysis is shaded on the X-axis. The peak of the N290 and P400 components is indicated with the arrows. The positioning of the electrodes included in the occipital cluster are shown in the electrode montage (see box and shaded electrode sites).

## Grand Average Waveforms by Stimulus Type at Anterior Temporal Electrodes

**Figure 3.**

The Auditory P1 component is shown for the unimodal visual (thin line), synchronous audiovisual (bold line), and asynchronous audiovisual (dashed line) conditions from Experiment 1. The Y-axis represents the amplitude of the ERP in microvolts, and the X-axis represents time following stimulus onset. The time-window of the component analysis is shaded on the X-axis. The positioning of the electrodes included in the anterior temporal clusters are shown within each electrode montage (see box and shaded electrode sites).

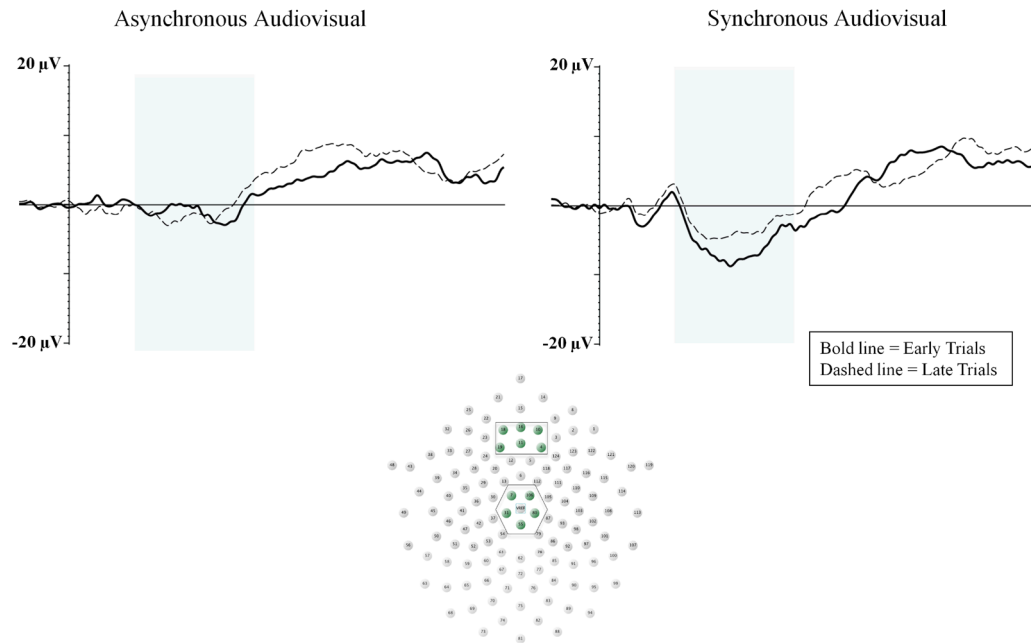
LSW by Stimulus Type and Trial Block at Temporal Electrodes



**Figure 4.** The LSW is shown for early (bold line) and late (dashed line) blocks of trials at temporal electrodes from Experiment 2. The left panel represents waveforms from the left hemisphere and the right panel represents waveforms from the right hemisphere. Waveforms from the synchronous audiovisual condition are shown on the top row, and waveforms from the asynchronous audiovisual condition are shown on the bottom row asynchronous audiovisual condition. The Y-axis represents the amplitude of the ERP in microvolts, and the X-axis represents time following stimulus onset. The time-window of the component analysis is shaded on the X-axis. The positioning of the electrodes included in the left and right temporal clusters are shown within the electrode montages (see boxes and shaded electrode sites).



## Nc ERP Component by Stimulus Type and Trial Block at Central Electrodes



**Figure 5.**

The Nc is shown for early (bold line) and late (dashed line) block of trials. The left panel represents the asynchronous audiovisual condition and the right panel represents the synchronous audiovisual condition from Experiment 2. The Y-axis represents the amplitude of the ERP in microvolts, and the X-axis represents time following stimulus onset. The time-window of the component analysis is shaded on the X-axis. The positioning of the electrodes included in the frontal and central midline clusters are shown within the electrode montage (see box and shaded electrode sites).