



Babies detect when the timing is right: Evidence from event-related potentials to a contingent mother-infant conversation

Christa Lam-Cassettari^{a,b,*}, Varghese Peter^{a,c}, Mark Antoniou^{a,b}

^a Western Sydney University, Australia

^b MARCS Institute for Brain, Behaviour and Development, Australia

^c School of Psychology, Western Sydney University, Australia

ARTICLE INFO

Keywords:

Event-related potentials (ERPs)
EEG
Speech perception
Mother-infant interaction
Contingent speech
Language acquisition

ABSTRACT

Social interactions are vital for healthy brain development. Burgeoning behavioural evidence indicates that a caregiver who provides contingently timed vocal responses to infant vocalisations provides key support for early language development. Understanding how contingently timed vocal responses relate to neurodevelopment in early infancy is lacking. This study compares event-related potentials (ERPs) to contingent and non-contingently timed vocalisations in 6- and 9-month-old infants ($n = 36$), and adults ($n = 24$). ERPs were recorded from each age group while listening to a naturalistic 21-minute recording of a mother playing and conversing with her baby. At 6-months, infants showed a significant positive ERP response to contingent vocalisations by the mother and infant. At 9-months infants showed negative ERP response to the mother's contingent speech. Adults showed no differences in ERPs between contingent and non-contingent speech regardless of the talker. We interpret the increased positivity in response to contingent speech as suggesting that infants show an attentional response at 6-months, and the increased negativity at 9-months relates to lexical-semantic processing. Further work is necessary to confirm the development of distinct ERPs shown in response to natural speech.

1. Introduction

Social interactions provide infants with unique opportunities to learn how to detect cues that provide meaning to speech and participate in vocal exchanges with others (Bornstein et al., 2015; Diehl et al., 2004). By their first birthday, most infants show an effortless progression from babbling speech-like syllables to uttering their first intelligible words (Dale and Fenson, 1996; Kalashnikova et al., 2016). Behavioural approaches have shown that the timing of conversations play a key role in reinforcing infant vocal behaviours and thus the development of early language skills. By the time infants are a few months old, they can participate in proto-conversations with their caregivers, who treat infant vocalisations as “meaningful” vocal turns and provide space for vocal behaviours (Snow, 1977). Importantly, infants who experience responsive vocalisations from a caregiver toward their own speech related vocalisations show developmental advances in their vocal behaviours (Warlaumont et al., 2014). Increased exposure to infant-directed speech at 6-, 10- and 14-months is also associated with larger vocabularies at 18 (Ferjan Ramirez et al., 2020) and 24-months (Ramirez-Esparza et al.,

2017a, 2017b). Thus, access to conversational turn-taking appears foundational for infant language development. But the question remains, how do infants come to know the “rules” for successful conversation? Can the dynamic nature of contingent parent-child conversations activate the neurodevelopment necessary for language processing? Designing an electrophysiological study that measures neural responses to a set of mother-infant conversations can provide a better understanding of how conversational timing impacts language processing.

It is known that infants undergo rapid neural development in the first 1000 days of life and that the experience that infants have with their environment modifies the structures and functions of the developing neural pathways (Fox et al., 2010). Thus, the quality of early social interaction has enduring effects on child language learning and processing. Child development research has witnessed a surge in neurophysiological studies attempting to delineate how infants come to understand and produce language by adapting measures used in adult studies. Event related potentials (ERPs) have been used to study how neural responses to speech change with infant age and provide evidence

* Corresponding author at: Western Sydney University, MARCS Institute for Brain, Behaviour and Development, Bankstown Campus, Locked Bag 1797, Penrith, NSW, 2751, Australia.

E-mail address: c.lam-cassettari@westernsydney.edu.au (C. Lam-Cassettari).

<https://doi.org/10.1016/j.dcn.2021.100923>

Received 10 December 2019; Received in revised form 23 December 2020; Accepted 8 January 2021

Available online 21 January 2021

1878-9293/© 2021 The Authors.

Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

for the cognitive processes involved in processing speech stimuli. ERPs reflect the pattern of electrical activity generated by large groups of brain cells in response to a sensory or cognitive event (Luck, 2014). Auditory ERPs to sound onsets in adult populations are typically shown by an N1-P2 response, with a negative peak around 90–110 ms and positive peak around 140–170 ms (Peter et al., 2016). In contrast, infants under 1 year of age show auditory ERPs to sound onsets as a P1-N2 response, with a positive peak around 150 ms and a negative peak around 250 ms (Peter et al., 2016).

Different ERP components have been used to assess stages of language processing. For example mismatch negativity/mismatch response (MMN/MMR) has been used to test infant discrimination of normal and rotated vowel stimuli between 4–8 months of age (Marklund et al., 2019). Closure positive shift (CPS) has been used to assess prosodic boundary perception in 2- to 6-year-old children (Männel et al., 2013; Männel and Friederici, 2009, 2011, 2016). The N400 has been used to study the lexico-semantic processing in 12 to 14-month-old children (Friedrich and Friederici, 2008, 2010). ERP effects have also been reported to show changes in sensitivity to mispronunciation in a consonant and vowel segmentation task at 8 months (Von Holzen et al., 2018), test the effect of word familiarity in familiarisation-test paradigms (Parise et al., 2010), and to examine the effect of prosodic accentuation on attention to speech (Männel and Friederici, 2013). The P3/P300 sub-components (P3a in non-attentive paradigms and P3b in attentive paradigms) provide another candidate for assessing developmental differences in the discrimination of auditory stimuli (McIsaac and Polich, 1992). Early ERP studies have shown that the P300 component is demonstrated in the same central-parietal distribution in 5- to 10-month-old infants and adults, albeit with a smaller amplitude and longer peak latency for infant groups when tested in a passive tone sequence paradigm (e.g., McIsaac and Polich, 1992). However, expression of the P300 appears to show greater variation in infancy and in paradigms testing responses to novel and unexpected stimuli (Riggins and Scott, 2020).

ERPs have high temporal resolution and many ERP components can be elicited without overt behavioural response, and are thus ideal for studying speech perception in early infancy. Electrophysiological studies in the first half year of life concur with robust behavioural evidence that the highly intonated infant-directed speech style is key to increasing infant attention to speech. At 5-months, exposure to familiar words such as the infant's own name, results in an N200-600 effect in the anterior region when compared to conditions in which infants hear an unfamiliar name (Parise et al., 2010). At 6 and 13 months, infants show a larger N600-800 in response to familiar words that are presented in infant-directed speech with heightened positive affect compared to the more monotone adult-directed speech style (Zangl and Mills, 2007). At 9-months, infants show evidence of semantic priming and exhibit the N400 component in the parietal area when they hear words spoken by their mother compared to an experimenter (Parise and Csibra, 2012). Together, these studies suggest that access to positive and highly intonated speech, and the repetitive infant-directed speech style supports language processing by activating both behavioural (e.g., Werker and McLeod, 1989) and neural attentional responses. Interestingly, stimuli that elicit different levels of emotional arousal in infants generate different ERP components. Eight-month-old infants show increased attention allocation and processing towards recorded infant vocalisations when they express positive compared to negative affect (Missana et al., 2017). When 8-month-old infants listen to recordings of infant cry vocalisations, ERPs showed an enhanced negativity around 200 ms at temporal electrodes; conversely, listening to recordings of infant laughter resulted in enhanced positivity around 300 ms at central electrodes (Missana et al., 2017).

Age-related changes in ERPs in the second half of the first year have been investigated using a familiarisation-test paradigm comparing 6-, 9- and 12-month-old infants' neural responses to words that were accented or words that were repeated (Männel and Friederici, 2013). ERPs

revealed differences in infants' word recognition as a function of situational input cues. By 6-months, prosodic accentuation elicited a positive response to familiarised words around 500 ms post word-onset. At 9-months, infants showed increased negativity to familiarised words around 400 ms post word-onset, while at 12-months, only word repetition led to increased negativity in speech processing. These developmental changes were hypothesised to be the result of both linguistic experience and maturation of the auditory cortex between 6–12 months. Studies assessing word segmentation in 10-month-old infants have reported a negativity over the left hemisphere at the onset of words as a marker of word detection/speech segmentation (Kooijman et al., 2013). At 7-months of age only a subset of infants showed this word-onset related negativity and those infants had superior language skills at 3 years of age compared to infants who did not show the ERP response at 7-months of age (Kooijman et al., 2013). At 9-months, infants exposed to a single vowel in an oddball paradigm show adult-like MMN responses to vowels spoken in infant-directed speech, but not adult-directed speech (Peter et al., 2016). By adulthood, vowels presented in both infant-directed and adult-directed speech generated the discriminatory MMN response (Peter et al., 2016). Given that infant ERP studies are typically modifications of adult paradigms, inclusion of an adult sample provides a language proficient baseline for developmental studies.

Accumulating neural evidence for infant speech processing abilities in the first year of life indicate that infants are sensitive to fine-grained differences in speech at the neural level, long before they can utter their first words. Furthermore, studies have shown variation in ERPs between approximately 6–9 months of age, arguably due to maturation of the auditory cortex in the second half of the first year of life. Studies to date have primarily compared speech that can be highly controlled and tested at the level of a single vowel or word. However, these studies do not test the association between the dynamic timing of spontaneous mother-infant vocal exchanges and neurodevelopmental advances for language processing.

Mothers who respond with well-timed social cues such as smiling, gesturing and vocalising have infants who produce more vocal behaviours (see Leclère et al., 2014 for a review). Infants who are immersed in a home environment rich in language are faster at processing familiar words in a looking-while-listening task, when infants hear a sentence containing an object name (e.g., dog) and must then look at a target image when it is pitted against a distractor (e.g., images of a dog vs. a baby) (Weisleder and Fernald, 2013). Importantly, infants that are engaged in more conversational turns in daily interactions have larger expressive vocabularies than peers who hear less speech and are involved in fewer conversations (Dwyer et al., 2018). Behavioural research with prelinguistic infants has shown that the quality and quantity of vocal attempts made by infants is correlated with the type of social feedback provided by the caregiver (Bornstein et al., 2015; Ferjan Ramirez et al., 2020; Goldstein and Schwade, 2008; Gros-Louis et al., 2006; Gros-Louis and Miller, 2018; Ramirez-Esparza et al., 2017b; Romeo et al., 2018; Tamis-LeMonda et al., 2001). Specifically, infants whose caregivers provided vocal feedback that was contingently responsive (i.e., timed) to their vocal behaviours, produced vocalisations that increasingly resembled the structure of their mother's vocalisations (Goldstein and Schwade, 2008). Conversely, infants in a yoked feedback condition did not produce babble that resembled the mother's vocal feedback to infant vocalisations (Goldstein and Schwade, 2008). More importantly, when caregivers were instructed to respond to their infant with contingent or non-contingent timed responses with respect to their infants' vocalisations, prelinguistic infants learnt new vocal forms only when mothers provided contingent speech to their vocalisations. Nine-month-old infants produced more mature vocalisations when their mother responded contingently to their babbling, compared to mothers who showed delayed responses (Goldstein and Schwade, 2008). Together these studies suggest that contingent responsiveness supports the language acquisition process.

Successful social interaction requires an understanding of the

dynamic back and forth nature of conversations, and the level of contingent responsiveness that will facilitate the positive progression of the interaction. Evidence suggests that contingent vocal responses that occur within 2 s of the speaker's last word provide optimal support for social interactions (Van Egeren et al., 2001). An international study by Bornstein et al. (2015) compared the similarities and differences in mother and infant vocal contingencies, and the overall frequency of mother-infant speech with 5.5-month-old infants in 11 countries (Argentina, Belgium, Brazil, Cameroon, France, Israel, Italy, Japan, Kenya, South Korea, United States). Interestingly, although the overall rate of mother and infant vocalisations was uncorrelated, mothers played a vital role in establishing turn-taking behaviours by providing contingent vocal responses to their infants' nondistress vocalisations. Moreover, infant nondistress vocalisations were contingent on their mothers' vocalisations. Despite consistent behavioural evidence that mothers' contingent vocal responsiveness enhances early speech production (Bornstein et al., 2015; Ferjan Ramírez et al., 2020; Goldstein and Schwade, 2008; Gros-Louis et al., 2006; McGillion et al., 2017; Romeo et al., 2018; Tamis-LeMonda et al., 2001), neurodevelopmental evidence is lacking. Understanding the neural mechanisms that support infant participation in dynamic interactions with their caregivers is critical to understanding how features that are readily available in the language learning environment support effective speech processing.

This study aims to determine prelinguistic (6- and 9-month-old) infants' ERPs to an audio-recording containing a series of naturally occurring contingent and non-contingent vocal turns between a mother and infant. The study was exploratory and therefore we adopted a data driven approach. No previous study has investigated the ERP effects to naturally occurring contingent compared to non-contingent mother-infant conversations. However, we were able to generate predictions about possible ERP effects by extrapolating from previous behavioural studies.

Behavioural studies have shown that infants whose parents produce more contingent vocalisations have larger vocabularies (e.g., Ferjan Ramírez et al., 2020; Tamis-LeMonda et al., 2001) and produce more sophisticated babble (Goldstein et al., 2003; Goldstein and Schwade, 2008) than peers who experience less contingent vocal feedback. One possible mechanism underlying this effect is that infants show earlier attentional responses to contingent vocalisations, and this heightened attention supports neurodevelopment in the first year of life. If this is the case, we might expect a response similar to the classic P3a response for contingent vocal stimuli (P3a is a positive response around 300 ms) that indicates an involuntary attention shift to the stimulus (Polich, 2012). A second possible mechanism is the ease of lexical access (how the words are accessed from the mental lexicon). The lexical access hypothesis can be indexed using the N400 response. It may be the case that infants generate an attentional response (similar to P3a) to contingent speech, or they may generate an N400-like response which has been shown in studies with infants as young as 9-months using a different paradigm (word-to-object priming) to the current study (see Parise and Csibra, 2012). However, the exploratory nature of this study does not support prediction of one component over the other. Given that pre-verbal infants typically experience an auditory environment rich in infant-directed speech; we expect infants to show a larger effect in response to the mother's vocal stimuli compared to the infant's vocal stimuli.

Adults are proficient language users, thus were included in this study to provide a reference sample to examine whether there are qualitative or quantitative differences in infant ERPs to pre-recorded contingent and non-contingent conversations. Since infant-directed speech contains more semantic information than an infant's vocalisations, adult participants were also expected to show larger language-related ERPs to the adult compared to infant speech stimuli. If alternatively, responses are based purely on the contingency timing and not the speaker type, adults should not show any difference in ERPs to adult versus infant speech.

2. Method

2.1. Ethics statement

The ethics committee for human research at Western Sydney University approved the experimental methods used in the study (Approval number: 11153). Written informed consent was obtained from adult participants and all mothers consented prior to participating in the study with their infant.

2.2. Participants

Power analysis was conducted in G*Power version 3.1 (Faul et al., 2007). Without previously published studies to refer to, we calculated our sample using a medium effect size (Cohen's $f = 0.25$) with a power of 0.80 at an alpha level of 0.05 for both within- and between-subject effects (3 groups, 4 conditions). Thus, a minimum of 45 subjects were required in total ($n = 15$ in each group). Attrition rates of 50 % are common in infant EEG studies (Stets et al., 2012), thus, we set out to test 30 subjects per age group to account for attrition and obtain 15 participants per age group, which is comparable to previous studies (e.g., Forgács et al., 2020; Quinn et al., 2006; Rivera-Gaxiola et al., 2005).

2.3. Adults

Twenty-four English speaking adults (17 females; M age: 18.5 years; range: 18–34 years) were recruited from the Western Sydney University undergraduate research participation system in exchange for course credit. All participants were native speakers of Australian English, reported normal hearing sensitivity, and were right-handed (one left handed) and were not parents or caregivers for young children. Adults did not report using infant-directed speech and had little to no experience interacting with infants. Data from one additional participant was rejected due to an insufficient number of artifact free trials (< 35 per condition).

2.4. Infants

Eighteen 6-month-old infants (7 females; age range: 5.9–7.3 months) and eighteen 9-month-old infants (8 females; age range: 8.8–9.9 months) participated in the study. Infants were recruited from MARCS Institute Babylab at Western Sydney University. All infant participants were being raised in an Australian English-speaking environment, born full-term, had normal hearing and no known developmental or neurological impairments. Participants were given a small gift (e.g., book, t-shirt) and parents received \$30 travel compensation. Twenty-three additional infants were tested but excluded from analysis due to an insufficient number of artifact free trials ($n = 4$ at 6 months; $n = 5$ at 9 months), technical faults ($n = 3$ at 6-months; $n = 3$ at 9 months), or fussiness/distress wearing the EEG cap leading to cessation of testing ($n = 4$ at 6 months; $n = 4$ at 9 months). The number of accepted infant participants were similar to previous ERP studies on language processing with infant participants of similar age (Forgács et al., 2020; Kabdebon et al., 2015; Kaduk et al., 2016; Quinn et al., 2006; Rivera-Gaxiola et al., 2005).

2.5. Stimuli

The audio recording used as stimuli for all participants in the current study was collected from a native Australian English-speaking mother and her 9-month-old son while they played (with a selection of age-appropriate soft toys and books) on a play mat in a sound attenuated audio booth. A natural infant-adult conversation was selected to increase the ecological validity of the stimulus; infants are typically immersed in speech input by their caregivers and partake in proto-conversations, and thus were expected to find this stimulus engaging.

The mother was instructed to play with her son as she normally would at home whilst wearing an EDIROL lapel microphone that was secured to her clothing prior to commencing the play session. The mother and infant were not given any instruction on how to talk during the play session to ensure that the recording contained spontaneous vocalisations. The mother spoke in infant-directed speech for the duration of the recording. A research assistant unaware of the study design and goals processed the audio recording to remove any non-speech vocalisations including heavy breathing, lip smacking or infant distress cries. Occasional silences greater than 5000 ms in length were trimmed using the following criteria to maintain integrity in the naturally random intervocalic timing (silence was reduced to 2500 ms + the duration of the utterance that preceded the prolonged period of silence e.g., an utterance of 482 ms followed by a 14000 ms silence, would be reduced to a silence 2500 ms + preceding utterance length 482 ms = 2982 ms). Prolonged segments of silence were removed from the stimuli recording to reduce the likelihood that participants would become fussy and lose interest in completing the experiment. Recorded utterances were classified as *contingent vocalisations* (speech that occurred within 1500 ms of the offset of the interactive partner's vocalisation) and *non-contingent vocalisations* (occurred > 2000 ms following the offset of the interactive partner's vocalisation). All stimuli were normalised to the same intensity level. Details of the stimuli count, acoustic measures and silence prior to trial onset for the contingent and non-contingent vocalisations are shown in Table 1. There were no differences in fundamental frequency between contingent and non-contingent vocalisations ($p > .05$), no difference between silence duration prior to the onset of contingent infant-other vocal turns ($p > .05$), or silence prior to the onset of non-contingent infant-mother vocal turns ($p > .05$). An excerpt of the recording is available in supplementary materials (Appendix A).

2.6. Procedure

EEG recording was conducted in a sound-attenuated, electrically-shielded room in a single block experiment lasting 21 min. While seated on their caregiver's lap, infants listened passively to an audio recording of a mother-infant play session presented through loudspeakers (model: Roland EDIROL MA:10A) placed on a desk 60 cm from the infant, at an RMS intensity of 70 dB SPL. Infants were shown an age-appropriate silent animation video and/or watched their caregiver play silently with puppets to minimise movement and fussiness and maintain infant engagement with the auditory stimuli. Adult participants listened to the same 21-minute audio recording while watching a silent movie of their choice (with subtitles). Adult participants were instructed to press a response button if they heard the word 'sheep' to maintain their attention to the auditory stimulus. In the audio file, triggers were placed at the target positions (i.e., at the onset of the speakers' contingent and non-contingent conversational turns). The stimulus presentation was controlled by Presentation 20.1 (Neurobehavioral Systems Inc, CA, USA), running on a PC. Triggers indicating the start of the contingent and non-contingent conversational turns were sent by Presentation and were included in the EEG recording to ensure synchronisation.

Continuous EEG was recorded using a 129 channel Hydrocel Geodesic Sensor Net (HCGSN), NetAmps 300 amplifier, and NetStation

5.1.2 software (EGI Inc.). EEG was recorded at a sampling rate of 1000 Hz. The reference electrode was placed at Cz. Impedances at electrode sites were kept below 50 k Ω . The test session took approximately 50 min (including net preparation, placement and removal).

2.7. EEG analysis

2.7.1. Adults

The offline analysis of the EEG was performed using Fieldtrip toolbox (v.20190404; Oostenveld et al., 2011) in MATLAB 2019a (Mathworks, Natick, MA, USA). The EEG signal was first bandpass filtered between 0.1–30 Hz using a windowed sinc finite impulse response filter. Epochs were created at the onset of each contingent or non-contingent conversational turn spoken by the mother or infant. It was then divided into epochs between -100 to 600 ms relative to the trigger indicating the onset of each contingent or non-contingent conversational turn (contingent adult, non-contingent adult, contingent infant, non-contingent infant). Epochs were baseline corrected between -100 to 0 ms. Independent component analysis (ICA) was then performed on the epoched data. Components with stereotypical features of eye blinks and eye movements were then removed from the data. Trials with absolute amplitude values exceeding $\pm 100 \mu\text{V}$ were removed. All participants had at least 35 trials accepted per condition. To ensure that there were no systematic signal to noise ratio differences between the contingent and non-contingent ERP responses (the number of presented trials were not equal across the conditions, Table 1), the same number of trials were entered into the analysis. Specifically, for each subject we identified the condition with the lowest number of artifact-free trials, and for other conditions we randomly selected the same number of trials from the artifact-free trials. The epochs were then re-referenced to the average of both mastoids. Trials were grand-averaged to generate the 4 ERP waves to be entered into the analysis for each participant: contingent adult, non-contingent adult, contingent infant, and non-contingent infant.

2.7.2. Infants

The infant EEG data were analysed as follows. To accommodate for the low frequency nature of infant responses (He et al., 2009), the data were bandpass filtered between 0.3–20 Hz. The continuous EEG was then divided into epochs from -100 to 600 ms post-stimulus onset. Baseline correction was performed between -100 to 0 ms. Consistent with common practice (Brusini et al., 2016; Kabdebon et al., 2015), channels located on the periphery of the sensor net were removed (36 channels) as noisy data in infants is consistently evident in these channels. Trials contaminated by artifacts were rejected on a trial-by-trial basis using a $\pm 100 \mu\text{V}$ criterion. Trials with more than 20 bad channels were removed. For trials with fewer than 20 bad channels, the channels with amplitude exceeding $\pm 100 \mu\text{V}$ were interpolated. Channels that were contaminated for more than 50 % of the trials were also interpolated. An inclusion criterion of >35 trials per condition was set a priori as previous studies have found language-related ERP components with a similar number of accepted trials (Hoehl and Wahl, 2012 e.g., p196). An equal number of clean trials were selected for all conditions and re-referenced to average mastoids, similar to the analysis on the adult data. Epochs were averaged separately for contingent and non-contingent conditions to create the ERP waveforms.

Table 1

Summary measures of the contingent and non-contingent stimuli by speaker vocalisation type, with fundamental frequency (F0) in Hertz, the period of silence (seconds:ms) and standard deviation of silence before trial onset markers began.

Speaker vocalisation type	Count	Mean F0	Minimum F0	Maximum F0	Minimum to maximum period of silence before trial onset (seconds:ms)	SD silence before trial onset (ms)
Infant Contingent	89	312	238	392	00:17 – 01:49	400
Infant Non-contingent	61	286	201	385	02:01 – 04:88	700
Mother Contingent	96	285	168	416	00:20 – 01:48	340
Mother Non-contingent	75	247	205	386	02:04 – 04:98	740

2.8. Statistical analysis

The initial analysis examined whether the ERPs for contingent and non-contingent conversations differed significantly from each other. If all electrodes and time points are included in the comparison, there will be 55,800 comparisons (93 electrodes \times 600 time points) and the chance of Type I error would be high. A classic approach in overcoming this is to limit the analysis to the electrodes and time points that have consistently showed an effect in the literature. However, this method has two main disadvantages: (1) it limits the analysis only to known effects; and (2) for the less studied effects like the contingency effect examined here, which is supported by behavioural research, there is not enough literature to identify the time points and electrodes for analysis. To overcome this, cluster-based mass permutation tests (Maris and Oostenveld, 2007) were computed separately for adult and infant vocalisations. This analysis is data driven and included all electrodes and time points in the epochs. First, a series of t -tests were computed for each electrode and time point. From this, time points where the difference between the waveforms was found to be significant ($p < .05$, two-tailed) were identified for each channel. Clusters were then formed based on temporal and spatial adjacency (at least 2 nearby time points and 2 nearby channels should show an effect) and on the polarity of the effect (all t values associated with a cluster should be either all positive or all negative). Cluster level statistics were computed by adding together all t values within a cluster (mass t score). To control Type I errors due to multiple comparisons, a permutation approach was used. For this, a data driven null hypothesis distribution was created by randomly swapping the stimuli labels within participants 2,000 times and computing mass t scores for each randomisation. The mass t scores obtained in the first step were then compared with the null hypothesis distribution. The cluster was determined to be significant if it fell in the top 2.5 or bottom 2.5 percentile of the null hypothesis distribution.

3. Results

3.1. Adults

The grand averaged ERP waveforms from 10 fronto-central electrodes (electrodes E4, E5, E6, E11, E12, E13, E19, E20, E112, E118 in HCGSN; for the location of these electrodes on the scalp, please see the supplementary figure in Appendix B) for contingent and non-contingent stimuli are shown in Fig. 1. As expected for auditory stimulation, a negative N1 peak was seen around 100 ms and a positive P2 peak around 200 ms. However, the cluster-based permutation test did not show any significant differences between the ERPs to contingent and non-contingent stimuli either for adult or infant vocalisations, suggesting that adults processed both contingent and non-contingent vocalisations

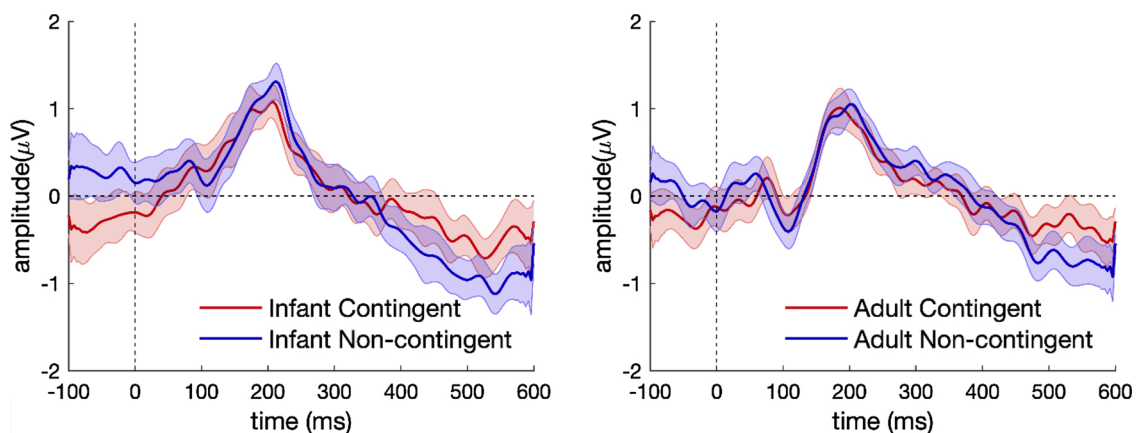


Fig. 1. ERPs did not differ to contingent and non-contingent speech stimuli in adult listeners. The dark lines show grand averaged ERPs and the shading encompasses 95 % confidence intervals.

similarly. Although the cluster-based permutation test included all the electrodes and all time points, the average response from 10 fronto-central electrodes are shown in Fig. 1 since no significant effects were identified in the analysis and a more relevant group of electrodes could not be identified. The auditory onset responses are more prominent in the fronto-central electrodes (Praat, 2011). ERPs from all electrodes are shown in Supplementary Appendix C.

3.2. Six-month-old infants

ERPs time locked to the onset of the contingent and non-contingent vocalisations in centro-parietal electrodes are shown in Fig. 2. The 6-month-old infants showed more positive ERPs to the contingent vocalisations in the 50–200 ms time window as compared to the non-contingent vocalisations. This effect was confirmed by the cluster-based permutation statistics (Table 2). Within the statistically significant cluster for the adult contingent versus non-contingent comparison, the contingent speech showed a more positive amplitude within the cluster ($M = 1.29 \mu\text{V}$, $SD = 2.42$) than the non-contingent speech ($M = -1.44 \mu\text{V}$, $SD = 3.76$). Similarly, for the infant vocalisations, the contingent vocalisations ($M = 0.65 \mu\text{V}$, $SD = 1.66$) showed a more positive amplitude than infant non-contingent vocalisations ($M = -1.35 \mu\text{V}$, $SD = 2.82$). The effect was mainly found at the centro-parietal electrodes (Fig. 2). The ERPs from all electrodes show a prominent P1 response in the fronto-central electrodes for all stimuli as expected for infants at this age (see Supplementary material).

To test whether statistical differences in contingent versus non-contingent effect was modulated by the speaker (infant vs. adult), we calculated the mean amplitude in a 40 ms time window around the peak of the effect (i.e., the peak at the contingent-non contingent difference waveform) in the grand averaged contingent-non contingent difference waveform from 10 electrodes in the centro-parietal region (Electrodes E31, E53, E54, E55, E61, E62, E78, E79, E80, E86 in HCGSN; Fig. 3). These electrodes were selected as they were part of the statistically significant cluster. The mean amplitude was calculated (as compared to peak amplitude) as it is less susceptible to noise (Luck, 2014) and we chose a 40 ms time window (98–138 ms in the infant vocalisation difference wave; 108–148 ms in the adult vocalisation difference wave) as the effect in the cluster based permutation test was in a relatively narrow time window (Table 2). Moreover, the mean amplitude does not change dramatically when the measurement window is altered (Luck, 2014). A paired sample t -test on the amplitude values did not show any significant effect, (infant contingent - infant non-contingent ($M = 0.90 \mu\text{V}$, $SE = 0.36$); adult contingent- adult non-contingent ($M = 1.05 \mu\text{V}$, $SE = 0.52$, $t(17) = -0.23$, $p > .05$, Cohen's d : -0.05) suggesting the effect was similar for both infant and adult vocalisations.

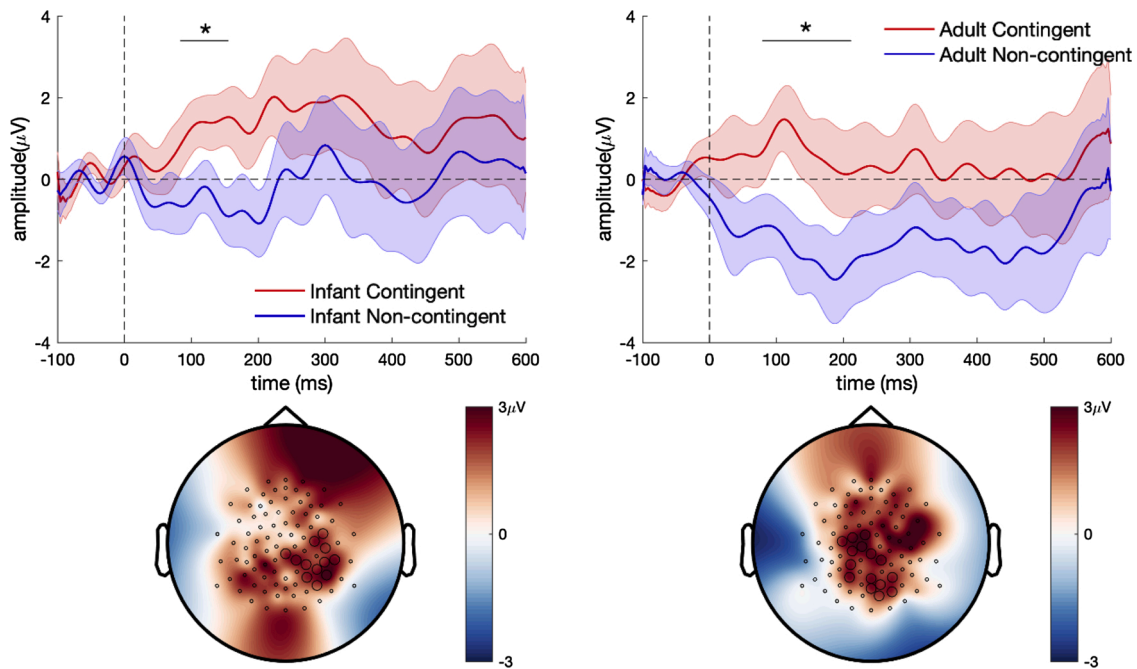


Fig. 2. ERPs from the 6-month-old infants to contingent and non-contingent speech stimuli (top panel) averaged from 10 centro-parietal electrodes. Topography of the contingent versus non-contingent effect is shown in the bottom panel. The circled electrodes belong to a statistically significant cluster.

Table 2
Results of the cluster-based permutation test on ERPs from 6-month-old infants.

Comparison	Cluster type	Time range	Cluster <i>p</i> value	Cohen's <i>d</i>
Adult contingent vs. non contingent vocalisations	Positive	80–212 ms	.041	0.88
Infant contingent vs. non contingent vocalisations	Positive	84–156 ms	.025	0.96

cluster-based permutation statistics showed a significant effect of contingency for adult vocalisations. The contingent vocalisations generated significantly more negative response clusters ($M = 3.85 \mu V, SD = 3.1$) as compared to non-contingent vocalisations ($M = 7.36 \mu V, SD = 4.3$) between 304–416 ms ($p = 0.039$, Cohen's $d = -0.89$). The cluster-based permutation statistics on infant vocalisations did not show any significant effects.

4. Discussion

This study used ERPs to examine infant and adult neural processing of a pre-recorded mother-infant play session containing authentic contingent and non-contingent vocal turns between a mother and her infant. The main findings showed that the 6-month-old group showed a larger positive response to contingent speech at the centro-parietal electrodes approximately between 50–200 ms compared to non-contingent speech; the 9-month-old group showed an enhanced negative response to adult contingent speech between 300–400 ms at fronto-central electrodes and an enhanced positivity to non-contingent stimuli; and adults showed no difference in neural processing of maternal speech or infant babbling regardless of whether it was a contingent or non-contingent vocal turn. This provides new evidence for distinct neural responses when infant listeners overhear audio-recordings containing contingent vocal turns between a mother and her infant. Notably, the results suggest that a passive listening task using a naturally recorded conversation may begin to evoke a semantic response before infants produce their first words.

The topography of the positive response for the 6-month-olds is similar to the attention-related ERP component P3 previously shown by adults, indicating an involuntary attention shift to the sound (Polich, 2012). However, the positivity in the current study occurs earlier (100–200 ms) than the P3 effect in adults (~300 ms). A positive response at a similar time range, but more anterior location has been reported in 5-month-old infants' responses to their own name compared to a stranger's name, which was attributed to increased attention to the auditory stimuli (Parise et al., 2010). Similarly, enhanced positivity around 150 ms was reported in 4-month-old infants listening to their mother's voice compared to unfamiliar voices (Purhonen et al., 2005),

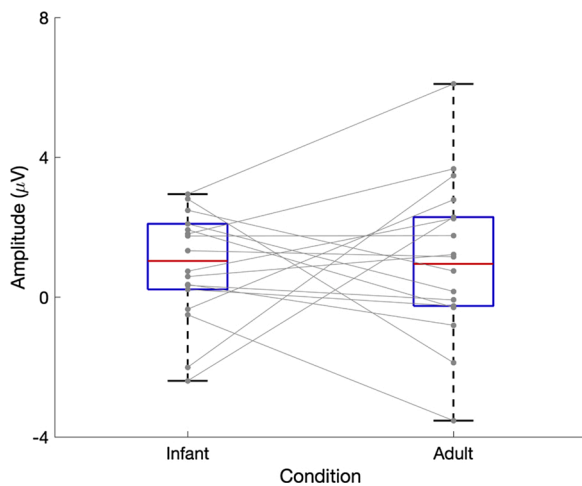


Fig. 3. Boxplots showing the mean amplitudes in the 40 ms time window around the positive peak in the contingent- non contingent difference waveform for 6-month-old infants to the infant and adult speech conditions.

3.3. Nine-month-old infants

The grand averaged ERP waveforms from fronto-central electrodes for contingent and non-contingent stimuli in 9-month-old infants is shown in Fig. 4. The ERPs showed a prominent P1 response for all stimuli at fronto-central electrodes (Supplementary material). The

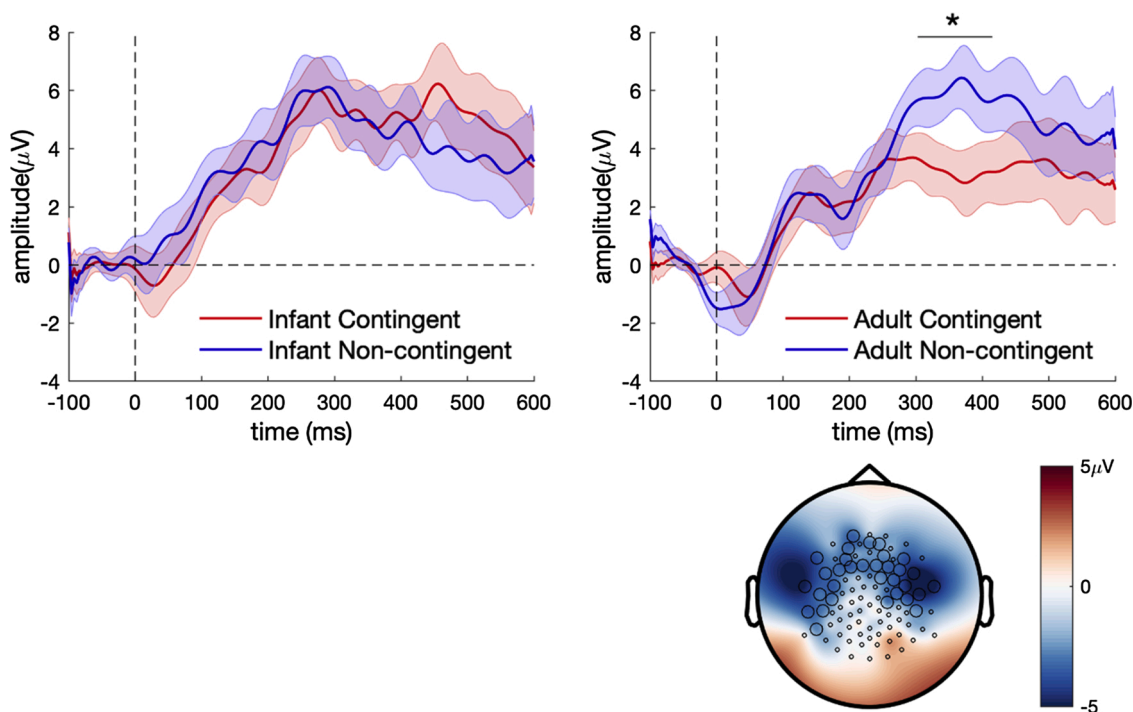


Fig. 4. ERPs from the 9-month-old infants to contingent and non-contingent speech stimuli (top panel). Topography of the contingent vs non-contingent effect is shown in the bottom panel. The circled electrodes belong to a statistically significant cluster.

which was interpreted as an indication of greater arousal due to hearing their mother's voice. Therefore, it is possible that the positivity in the 100–200 ms time range in 6-month-old infants reflects their increased attention and arousal to contingent over non-contingent speech, regardless of whether the speaker was a mother or an infant.

The ERP data for the 9-month-old group shows a negative ERP component in response to contingent adult speech, but not contingent infant vocalisations. Since the time range and the topography of the effect at 9-months of age is suggestive of the classic N400 effect (Kutas and Federmeier, 2011), this effect is interpreted as the N400 to contingent adult vocalisations. This provides preliminary evidence that at 9-months, when infants are showing attunement to their native language (Werker and Tees, 1984), progressing into the canonical babbling stage and producing more mature phonetic syllables (Oller, 1980), infants are beginning to process conversational turns in a semantic and lexical nature. Indeed, the ERP response at around 350 ms is reflective of semantic/lexical processing (Friederici, 2005). The N400 is typically recorded in paradigms involving semantic violation and it is believed to reflect semantic integration difficulties (for a review see Kutas and Federmeier, 2011). However, the N400 response is present while processing any conceptually meaningful stimuli (Szewczyk and Schriefers, 2018). The amplitude of N400 reflects the extent to which the meaning of the stimulus is congruent with the preceding context; congruent words generate a small N400 and incongruous words generate a large N400 response (Lau et al., 2008). According to the lexical pre-activation view (Federmeier, 2007; Kutas and Federmeier, 2000), the N400 is a by-product of lexical access. Lexical access is easy for congruent words as the words are activated to some degree by the preceding context, thereby generating N400 with a small amplitude. But when the word is incongruent (and not preactivated by the preceding context), lexical access is effortful, and thus a larger N400 is generated. In accord with this evidence, we propose that an N400 generated by adult contingent vocalisations for the group of 9-month-old infants is indicative of the lexical access process, although effortful at this stage. This line of reasoning raises the question whether there is any lexical access for the non-contingent adult vocalisations. A new study would be required

where both the contingency (contingent, non-contingent) and lexical nature (lexical, non-lexical) of adult vocalisations are manipulated to directly test whether there is indeed lexical access for the non-contingent adult vocalisations. We interpret the presence of the N400-like response in the data as suggesting that recordings of real-world auditory stimuli may elicit semantic processing earlier than shown by previous studies that have shown the N400 using word-priming paradigms with 12- to 14-month-old infants (Friedrich and Friederici, 2008, 2010).

We interpret the findings that infant contingent vocalisations did not generate N400 as indicating that the infant vocalisations were less meaningful in nature, as they were babbled phonetic syllables, and therefore no lexical access was performed. One recent study in adults using conversational speech showed an N400 effect while comparing the response to predictable versus unpredictable words (Rasenberg et al., 2020). However, the lexical congruency of the contingent and non-contingent stimuli were not explicitly manipulated in the current experiment.

A possible reason for the absence of any significant effect in adults (either the positivity or N400-like response) could be due to the task. Similarly to Becker et al. (2014), the absence of the N400 effect in adults may be related to their passive listening task. Adults were given a simple task (press a button when they heard the words 'shark', 'sheep' or 'shoe' on the audio recording because they were toys the mother-infant pair played with) to encourage their attention to the speech stimuli for the duration of the study, however, this task did not encourage attention to the contingent versus non-contingent timing of the vocal stimuli. Many ERP effects for linguistic stimuli are modulated by the task demands (Schacht et al., 2014), thus it is plausible that the absence of any effects in the adult group are due to the type of task that was used because it limited the amount of lexical processing. Furthermore, the stimuli were natural speech and could vary in terms of characteristics such as the mother's affective vocal tone (e.g., how comforting, approving or directive speech sounds (Kitamura and Burnham, 2003), which has been shown to elicit differential preferences from 3-, 6- and 9-month-old infants (Kitamura and Lam, 2009). Consequently, non-parent adult

listeners may have disengaged from listening to the overheard conversation as they lacked interest in the task. Future work in which adult participants are tasked with attending to the timing of the mother-infant conversation turns is required to test this hypothesis.

We argue that the N400-like response shown by the 9-month-old group is the result of the ecological validity of the study due to the naturally recorded mother-infant linguistic stimuli. Previous research on the N400 has shown that only infants with larger vocabularies show an N400 response at 12-months of age (Friedrich and Friederici, 2010) whereas all infants show this response at 14-months (Friedrich and Friederici, 2008). The N400 response obtained at an earlier age was only under specific stimulus conditions. Parise and Csibra (2012) showed an N400 response to picture-word mismatch in 9-month-old infants, only when the speaker was the infant's mother. Junge et al. (2012) showed N400 to picture-word mismatches in 9-month-old infants after extensive object-word familiarisation trials. All of these previous studies involving infants used a picture-word mismatch paradigm, whereas the present study showed an N400 response using a recording of natural speech alone. Future work examining neural responses to contingently timed non-speech sounds will help delineate whether these findings are limited to speech contingency or simply the nature of contingent timing.

An alternative explanation for the N400-type response to contingent stimuli is the possibility of a closure positive shift (CPS) response at the phrase boundaries to non-contingent stimuli (Steinhauer et al., 1999). The CPS response in preschool children is known to be sensitive to pause duration (Männel and Friederici, 2016). The non-contingent stimuli in the present study had a pause duration of 2000 ms or more whereas the contingent stimuli had a pause duration less than 1500 ms. Therefore, a potential explanation for our pattern of results is that the longer pause duration in the non-contingent stimuli could be responsible for generating the observed positive response, which we incorrectly interpreted as an N400 response to contingent stimuli. However, this explanation is not supported by the data because the N400 effect only emerged for non-contingent adult vocalisations and not those made by the infant. If the effect is due to low level acoustic properties such as pause duration, we would expect the effect for both adult and infant vocalisations and this was not evident in the data.

Increased parent-infant conversations are known to support positive language development in the first years of life. This study provides key insights into the effect of contingent conversations on the development of neural responses to speech, showing that the contingent timing of parent-infant conversations give rise to distinct neural responses to speech timing by 6-months of age, but do not affect the neural responses of language proficient adults. The infant findings concur with behavioural evidence which shows that caregivers who respond to infant vocal behaviours with contingently timed vocal responses encourage and strengthen the complexity of infant vocal behaviours in real-time (Bornstein et al., 2015; Goldstein and Schwade, 2008; Gros-Louis et al., 2006; Reed et al., 2017; Romeo et al., 2018; Tamis-LeMonda et al., 2014). Future research combining intervention approaches with electrophysiological methods are warranted to provide further insight into how exposure to contingent vocal responses shape language development in the first years of life.

Importantly, this study used naturally recorded stimuli from a mother-infant interaction, thus, the stimuli contain contingent and non-contingent timing characteristics, as well as the spontaneously exaggerated acoustic (Fernald and Simon, 1984) and affective (Kitamura and Burnham, 2003; Kitamura and Lam, 2009) qualities of the IDS register that are elicited while interacting with a child. No statistical difference was shown in the fundamental frequency or amplitude of the mother's or infant's contingent compared to non-contingent vocalisations. Thus, this study illuminates conversational timing as a mechanism that requires further investigation, particularly in populations of infants who receive less speech input from caregivers, such as infants of postnatally depressed mothers (Lam-Cassettari and Kohlhoff, 2020) or infants experiencing hearing loss (Lam-Cassettari et al., 2015). Further research

is needed to extend these findings. Our cross-sectional evidence suggests that neural responses change with maturation as the 6-month-old group showed distinctly attentional responses, however, the 9-month-old group showed qualitatively different ERPs. Given that adults showed no difference in neural responses to contingent and non-contingent speech, longitudinal follow-up in later infancy is necessary to determine precisely when infants respond to timing of mother-infant conversational turns semantically. Further prospective work should also assess the relationship between infant language outcomes and the level of exposure infants have with contingently responsive vocal interactions.

This study provides new insight into the neurodevelopmental responses of infants and adults listening to recordings of naturally occurring contingent and non-contingent conversations between a mother and her infant. Overall, the results provide evidence that the infant brain responds differently to contingent compared to non-contingent vocal turns between a mother and infant. Adults show no difference in brain potentials in response to contingent compared to non-contingent mother-infant conversations which we interpret as a consequence of the superficial lexical demand required by the task assigned to adults in the passive listening task. In this study infants showed distinct ERPs to contingent and non-contingent speech, thus providing preliminary ERP support to the growing body of behavioural evidence demonstrating that access to contingently responsive infant-directed conversations is critical to language development in the first years of life.

CRediT authorship contribution statement

Christa Lam-Cassettari: Conceptualization, Funding acquisition, Methodology, Project administration, Investigation, Supervision, Writing - original draft, Writing - review & editing. **Varghese Peter:** Conceptualization, Methodology, Supervision, Formal analysis, Writing - original draft, Writing - review & editing. **Mark Antoniou:** Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by an Australian Research Council Transdisciplinary Innovation Grant from the Centre of Excellence on the Dynamics of Language; and a Marcus and Amalia Wallenberg Foundation grant, Sweden (2013.0056) awarded to CLC. We thank all the participants (adults, infants and their families) who participated in this study, Brittany and Caitlin for assistance with data collection and the anonymous reviewers.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.dcn.2021.100923>.

References

- Becker, A.B.C., Schild, U., Friedrich, C.K., 2014. ERP correlates of word onset priming in infants and young children. *Dev. Cogn. Neurosci.* 9, 44–55. <https://doi.org/10.1016/j.dcn.2013.12.004>.
- Bornstein, M.H., Putnick, D.L., Cote, L.R., Haynes, O.M., Suwalsky, J.T.D., 2015. Mother-infant contingent vocalizations in 11 countries. *Psychol. Sci.* 26 (8), 1272–1284. <https://doi.org/10.1177/0956797615586796>.
- Brusini, P., Dehaene-Lambertz, G., Dutat, M., Goffinet, F., Christophe, A., 2016. ERP evidence for on-line syntactic computations in 2-year-olds. *Dev. Cogn. Neurosci.* 19, 164–173. <https://doi.org/10.1016/j.dcn.2016.02.009>.

- Dale, P.S., Fenson, L., 1996. Lexical development norms for young children. *Behav. Res. Methods Instrum. Comput.* 28 (1), 125–127. <https://doi.org/10.3758/BF03203646>.
- Diehl, R.L., Lotto, A.J., Holt, L.L., 2004. Speech perception. *Annu. Rev. Psychol.* 55 (1), 149–179. <https://doi.org/10.1146/annurev.psych.55.090902.142028>.
- Dwyer, A., Jones, C., Davis, C., Kitamura, C., Ching, T.Y.C., 2018. Maternal education influences Australian infants' language experience from six months. *Infancy* 24 (1), 90–100. <https://doi.org/10.1111/inf.12262>.
- Federmeier, K.D., 2007. Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44 (4), 491–505. <https://doi.org/10.1111/j.1469-8986.2007.00531.x>.
- Ferjan Ramírez, N., Lytle, S.R., Kuhl, P.K., 2020. Parent coaching increases conversational turns and advances infant language development. *Proc. Natl. Acad. Sci. U. S. A.* 117 (7), 3484–3491. <https://doi.org/10.1073/pnas.1921653117>.
- Fernald, A., Simon, T., 1984. Expanded intonation contours in mothers' speech to newborns. *Dev. Psychol.* 20 (1), 104–113. <https://doi.org/10.1037/0012-1649.20.1.104>.
- Forgács, B., Gervain, J., Parise, E., Csibra, G., Gergely, G., Baross, J., Király, I., 2020. Electrophysiological investigation of infants' understanding of understanding. *Dev. Cogn. Neurosci.* 43, 100783 <https://doi.org/10.1016/j.dcn.2020.100783>.
- Fox, S.E., Levitt, P., Nelson L.L.I., C.A., 2010. How the timing and quality of early experiences influence the development of brain architecture. *Child Dev.* 81 (1), 28–40. <https://doi.org/10.1111/j.1467-8624.2009.01380.x>.
- Friederici, A.D., 2005. Neurophysiological markers of early language acquisition: from syllables to sentences. *Trends Cogn. Sci.* 9 (10), 481–488. <https://doi.org/10.1016/j.tics.2005.08.008>.
- Friedrich, M., Friederici, A.D., 2008. Neurophysiological correlates of online word learning in 14-month-old infants. *NeuroReport* 19 (18), 1757–1761. <https://doi.org/10.1097/WNR.0b013e328318f014>.
- Friedrich, M., Friederici, A.D., 2010. Maturing brain mechanisms and developing behavioral language skills. *Special Issue on Language Development* 114 (2), 66–71. <https://doi.org/10.1016/j.bandl.2009.07.004>.
- Goldstein, M.H., Schwade, J.A., 2008. Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19 (5), 515–523. <https://doi.org/10.1111/j.1467-9280.2008.02117.x>.
- Goldstein, M.H., King, A.P., West, M.J., 2003. Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci. U. S. A.* 100 (13), 8030–8035. <https://doi.org/10.1073/pnas.1332441100>.
- Gros-Louis, J., Miller, J.L., 2018. From 'ah' to 'bah': social feedback loops for speech sounds at key points of developmental transition. *J. Child Lang.* 45 (3), 807–825. <https://doi.org/10.1017/S0305000917000472>.
- Gros-Louis, J., West, M.J., Goldstein, M.H., King, A.P., 2006. Mothers provide differential feedback to infants' prelinguistic sounds. *Int. J. Behav. Dev.* 30 (6), 509–516. <https://doi.org/10.1177/0165025406071914>.
- He, C., Hotson, L., Trainor, L.J., 2009. Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. *Eur. J. Neurosci.* 29 (4), 861–867. <https://doi.org/10.1111/j.1460-9568.2009.06625.x>.
- Hoehl, S., Wahl, S., 2012. Recording infant ERP data for cognitive research. *Dev. Neuropsychol.* 37 (3), 187–209. <https://doi.org/10.1080/87565641.2011.627958>.
- Junge, C., Cutler, A., Hagoort, P., 2012. Electrophysiological evidence of early word learning. *Neuropsychologia* 50 (14), 3702–3712. <https://doi.org/10.1016/j.neuropsychologia.2012.10.012>.
- Kabdebon, C., Pena, M., Buiatti, M., Dehaene-Lambertz, G., 2015. Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain Lang.* 148, 25–36. <https://doi.org/10.1016/j.bandl.2015.03.005>.
- Kaduk, K., Bakker, M., Juvrud, J., Gredebäck, G., Westermann, G., Lunn, J., Reid, V.M., 2016. Semantic processing of actions at 9 months is linked to language proficiency at 9 and 18 months. *J. Exp. Child Psychol.* 151, 96–108. <https://doi.org/10.1016/j.jecp.2016.02.003>.
- Kalashnikova, M., Schwarz, I.-C., Burnham, D., 2016. OZI: australian english communicative development inventory. *First Lang.* 36 (4), 407–427. <https://doi.org/10.1177/014273716648846>.
- Kitamura, C., Burnham, D., 2003. Pitch and communicative intent in mother's speech: adjustments for age and sex in the first year. *Infancy* 4 (1), 85–110. https://doi.org/10.1207/S15327078IN0401_5.
- Kitamura, C., Lam, C., 2009. Age-specific preferences for infant-directed affective intent. *Infancy* 14 (1), 77–100. <https://doi.org/10.1080/15250000802569777>.
- Kooijman, V., Junge, C., Johnson, E.K., Hagoort, P., Cutler, A., 2013. Predictive brain signals of linguistic development. *Front. Psychol.* 4 <https://doi.org/10.3389/fpsyg.2013.00025>.
- Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4 (12), 463–470. [https://doi.org/10.1016/S1364-6613\(00\)01560-6](https://doi.org/10.1016/S1364-6613(00)01560-6).
- Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62 (1), 621–647. <https://doi.org/10.1146/annurev.psych.093008.131123>.
- Lam-Cassettari, C., Kohlhoff, J., 2020. Effect of maternal depression on infant-directed speech to prelinguistic infants: implications for language development. *PLoS One*. <https://doi.org/10.1371/journal.pone.0236787>.
- Lam-Cassettari, C., Wadnerkar-Kamble, M.B., James, D.M., 2015. Enhancing parent-child communication and parental self-esteem with a video-feedback intervention: outcomes with prelingual deaf and hard-of-hearing children. *J. Deaf Stud. Deaf Educ.* 20 (3), 266–274. <https://doi.org/10.1093/deafed/env008>.
- Lau, E.F., Phillips, C., Poeppl, D., 2008. A cortical network for semantics: (De)constructing the N400. *Nat. Rev. Neurosci.* 9 (12), 920–933. <https://doi.org/10.1038/nrn2532>.
- Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., Cohen, D., 2014. Why synchrony matters during mother-child interactions: a systematic review. *PLoS One* 9 (12), e113571. <https://doi.org/10.1371/journal.pone.0113571>.
- Luck, S.J., 2014. *An Introduction to the Event-Related Potential Technique*, 2nd ed. MIT Press <https://mitpress.mit.edu/books/introduction-event-related-potential-technique-second-edition>.
- Männel, C., Friederici, A.D., 2009. Pauses and intonational phrasing: ERP studies in 5-month-old German infants and adults. *J. Cogn. Neurosci.* 21 (10), 1988–2006. <https://doi.org/10.1162/jocn.2009.21221>.
- Männel, C., Friederici, A.D., 2011. Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children: intonational phrase structure processing. *Dev. Sci.* 14 (4), 786–798. <https://doi.org/10.1111/j.1467-7687.2010.01025.x>.
- Männel, C., Friederici, A.D., 2013. Accentuate or repeat? Brain signatures of developmental periods in infant word recognition. *Cortex* 49 (10), 2788–2798. <https://doi.org/10.1016/j.cortex.2013.09.003>.
- Männel, C., Friederici, A.D., 2016. Neural correlates of prosodic boundary perception in German preschoolers: if pause is present, pitch can go. *Brain Res.* 1632, 27–33. <https://doi.org/10.1016/j.brainres.2015.12.009>.
- Männel, C., Schipke, C.S., Friederici, A.D., 2013. The role of pause as a prosodic boundary marker: language ERP studies in German 3- and 6-year-olds. *Dev. Cogn. Neurosci.* 5, 86–94. <https://doi.org/10.1016/j.dcn.2013.01.003>.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164 (1), 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Marklund, E., Schwarz, I.-C., Lacerda, F., 2019. Amount of speech exposure predicts vowel perception in four- to eight-month-olds. *Dev. Cogn. Neurosci.* 36, 100622 <https://doi.org/10.1016/j.dcn.2019.100622>.
- McGillion, M., Pine, J.M., Herbert, J.S., Matthews, D., 2017. A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *J. Child Psychol. Psychiatry* 58 (10), 1122–1131. <https://doi.org/10.1111/jcpp.12725>.
- Mclsaac, H., Polich, J., 1992. Comparison of infant and adult P300 from auditory stimuli. *J. Exp. Child Psychol.* 53 (2), 115–128. [https://doi.org/10.1016/0022-0965\(92\)90044-7](https://doi.org/10.1016/0022-0965(92)90044-7).
- Missana, M., Altvater-Mackensen, N., Grossmann, T., 2017. Neural correlates of infants' sensitivity to vocal expressions of peers. *Dev. Cogn. Neurosci.* 26, 39–44. <https://doi.org/10.1016/j.dcn.2017.04.003>.
- Oller, D.K., 1980. *Production*. In: *Child Phonology*, Vol 1. Academic Press, pp. 93–112.
- Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011, 1–9. <https://doi.org/10.1155/2011/156869>.
- Parise, E., Csibra, G., 2012. Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychol. Sci.* 23 (7), 728–733. <https://doi.org/10.1177/0956797612438734>.
- Parise, E., Friederici, A.D., Striano, T., 2010. "Did you call me?" 5-month-old infants own name guides their attention. *PLoS One* 5 (12), e14208. <https://doi.org/10.1371/journal.pone.0014208>.
- Peter, V., Kalashnikova, M., Santos, A., Burnham, D., 2016. Mature neural responses to infant-directed speech but not adult-directed speech in pre-verbal infants. *Sci. Rep.* 6 (1), 34273. <https://doi.org/10.1038/srep34273>.
- Polich, J., 2012. *Neuropsychology of P300*. The Oxford Handbook of Event-related Potential Components, Oxford.
- Praet, H., 2011. *Sensory ERP components*. In: Luck, Steven J., Kappenman, E.S. (Eds.), *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press, pp. 89–114.
- Purhonen, M., Kilpeläinen-Lees, R., Valkonen-Korhonen, M., Karhu, J., Lehtonen, J., 2005. Four-month-old infants process own mother's voice faster than unfamiliar voices—Electrical signs of sensitization in infant brain. *Cogn. Brain Res.* 24 (3), 627–633. <https://doi.org/10.1016/j.cogbrainres.2005.03.012>.
- Quinn, P.C., Westerlund, A., Nelson, C.A., 2006. Neural markers of categorization in 6-month-old infants. *Psychol. Sci.* 17 (1), 59–66. <https://doi.org/10.1111/j.1467-9280.2005.01665.x>.
- Ramírez-Esparza, N., García-Sierra, A., Kuhl, P.K., 2017a. Look who's talking NOW! Parentese speech, social context, and language development across time. *Front. Psychol.* 8 (JUNE) <https://doi.org/10.3389/fpsyg.2017.01008>. Scopus.
- Ramírez-Esparza, N., García-Sierra, A., Kuhl, P.K., 2017b. The impact of early social interactions on later language development in Spanish-english bilingual infants. *Child Dev.* 88 (4), 1216–1234. <https://doi.org/10.1111/cdev.12648>.
- Rasenberg, M., Rommers, J., van Bergen, G., 2020. Anticipating predictability: an ERP investigation of expectation-managing discourse markers in dialogue comprehension. *Lang. Cogn. Neurosci.* 35 (1), 1–16. <https://doi.org/10.1080/23273798.2019.1624789>.
- Reed, J., Hirsh-Pasek, K., Golinkoff, R.M., 2017. Learning on hold: cell phones sidetrack parent-child interactions. *Dev. Psychol.* 53 (8), 1428–1436. <https://doi.org/10.1037/dev0000292>.
- Riggins, T., Scott, L.S., 2020. P300 development from infancy to adolescence. *Psychophysiology* 57 (7). <https://doi.org/10.1111/psyp.13346>.
- Rivera-Gaxiola, M., Silva-Pereyra, J., Kuhl, P.K., 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev. Sci.* 8 (2), 162–172. <https://doi.org/10.1111/j.1467-7687.2005.00403.x>.
- Romeo, R.R., Leonard, J.A., Robinson, S.T., West, M.R., Mackey, A.P., Rowe, M.L., Gabrieli, J.D.E., 2018. Beyond the 30-million-word gap: children's conversational exposure is associated with language-related brain function. *Psychol. Sci.* 29 (5), 700–710. <https://doi.org/10.1177/0956797617742725>.

- Schacht, A., Sommer, W., Shmuilovich, O., Martíenz, P.C., Martín-Loeches, M., 2014. Differential task effects on N400 and P600 elicited by semantic and syntactic violations. *PLoS One* 9 (3), e91226. <https://doi.org/10.1371/journal.pone.0091226>.
- Snow, C.E., 1977. The development of conversation between mothers and babies *. *J. Child Lang.* 4 (1), 1–22. <https://doi.org/10.1017/S030500090000453>.
- Steinhauer, K., Alter, K., Friederici, A., 1999. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* 2, 191–196. <https://doi.org/10.1038/5757>.
- Stets, M., Stahl, D., Reid, V.M., 2012. A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Dev. Neuropsychol.* 37 (3), 226–252. <https://doi.org/10.1080/87565641.2012.654867>.
- Szewczyk, J.M., Schriefers, H., 2018. The N400 as an index of lexical preactivation and its implications for prediction in language comprehension. *Lang. Cogn. Neurosci.* 33 (6), 665–686. <https://doi.org/10.1080/23273798.2017.1401101>.
- Tamis-LeMonda, C.S., Bornstein, M.H., Baumwell, L., 2001. Maternal responsiveness and children's achievement of language milestones. *Child Dev.* 72 (3), 748–767. <https://doi.org/10.1111/1467-8624.00313>.
- Tamis-LeMonda, C.S., Kuchirko, Y., Song, L., 2014. Why Is infant language learning facilitated by parental responsiveness? *Curr. Dir. Psychol. Sci.* 23 (2), 121–126. <https://doi.org/10.1177/0963721414522813>.
- Van Egeren, L.A., Barratt, M.S., Roach, M.A., 2001. Mother–infant responsiveness: timing, mutual regulation, and interactional context. *Dev. Psychol.* 37 (5), 684–697. <https://doi.org/10.1037/0012-1649.37.5.684>.
- Von Holzen, K., Nishibayashi, L.-L., Nazzi, T., 2018. Consonant and vowel processing in word form segmentation: an infant ERP study. *Brain Sci.* 8 (2), 24. <https://doi.org/10.3390/brainsci8020024>.
- Warlaumont, A.S., Richards, J.A., Gilkerson, J., Oller, D.K., 2014. A social feedback loop for speech development and its reduction in autism. *Psychol. Sci.* 25 (7), 1314–1324. <https://doi.org/10.1177/0956797614531023>.
- Weisleder, A., Fernald, A., 2013. Talking to children matters: early language experience strengthens processing and builds vocabulary. *Psychol. Sci.* 24 (11), 2143. <https://doi.org/10.1177/0956797613488145>.
- Werker, J.F., McLeod, P., 1989. Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43 (2), 230–246.
- Werker, J.F., Tees, R.C., 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behav. Dev.* 7 (1), 49–63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3).
- Zangl, R., Mills, D.L., 2007. Increased brain activity to infant-directed speech in 6- and 13-Month-Old infants. *Infancy* 11 (1), 31–62. https://doi.org/10.1207/s15327078in1101_2.