

Published in final edited form as:

Psychol Sci. 2012 July 1; 23(7): 728–733. doi:10.1177/0956797612438734.

Electrophysiological Evidence for the Understanding of Maternal Speech by 9-Month-Old Infants

Eugenio Parise and **Gergely Csibra**

Cognitive Development Center, Department of Cognitive Science, Central European University, Budapest, Hungary

Abstract

Early word learning in infants relies on statistical, prosodic, and social cues that support speech segmentation and the attachment of meaning to words. It is debated whether such early word knowledge represents mere associations between sound patterns and visual object features, or reflects referential understanding of words. By using event-related brain potentials, we demonstrate that 9-month-old infants detect the mismatch between an object appearing from behind an occluder and a preceding label with which their mother introduces it. The N400 effect has been shown to reflect semantic priming in adults, and its absence in infants has been interpreted as a sign of associative word learning. By setting up a live communicative situation for referring to objects, we demonstrate that a similar priming effect also occurs in young infants. This finding may indicate that word meaning is referential from the outset, and it drives, rather than results from, vocabulary acquisition in humans.

Keywords

Infants; Language Acquisition; Semantic Priming; ERP; N400

Word learning in infancy is supported by a range of cognitive skills (Kuhl, 2004). By at least 8 months of age, infants are able to use statistical information to segment words from continuous speech (Saffran, Aslin, & Newport, 1996), and rely on prosody both in word segmentation and in associating novel words with visual referents (Shukla, White, & Aslin, 2011). Mapping words onto semantic representations by matching the word form to a referent is a non-trivial problem (Gleitman & Gleitman, 1992). There is no agreement whether young infants are capable of referential word understanding that is genuinely semantic (Waxman & Gelman, 2009). Evidence suggests that infants conceive deictic signals, such as gazing and pointing, as expressing communicative reference (e.g. Senju & Csibra, 2008), exploit such signals in word learning (e.g. Hollich, et al., 2000), and one-year-olds expect that concurrent verbal and non-verbal expressions from the same source co-refer to the same object, suggesting that they appreciate the referential nature of both pointing and words (Gliga & Csibra, 2009). However, some researchers have argued that early word understanding reflects simple associations between auditory and visual stimuli

without manifesting the understanding of referential and symbolic nature of words (Robinson, Howard, & Sloutsky, 2005).

Whether infants form stable word-object associations that reflect referential understanding of speech at this age is not yet known. One way to test such understanding in infants is using event-related brain potentials (ERPs). The N400 ERP component has been shown to reflect semantic priming in adults (Kutas & Hillyard, 1980), and Friedrich and Friederici (2004, 2005a, 2005b) found that adults, 19- and 14-month-olds, but not 12-month-olds, produced an enhanced N400 to words incongruous with picture primes. In a follow-up study, only 12-month-olds who were assessed as high word producers showed a N400 effect (Friedrich & Friederici, 2010). The absence of N400 effect in infants younger than 12 months was interpreted as a sign of merely associative word understanding, which does not entail semantic processing. Word learning requires the learner to form associations between word forms and visual referents. However, temporal associations alone do not imply semantic representation of word meaning. In the present study, we attempt to clarify *what* infants learn when they form word-object associations: meaningless links or semantically meaningful sign-referent relations.

In order to shed light whether young infants interpret words referentially and semantically, we developed a new paradigm that overcomes some shortcomings of earlier ERP studies with infants. In particular, we wanted to ensure that infants realize that the word they hear refers to the object they see. Since both theoretical arguments (Csibra, 2010) and empirical findings (e.g., Senju, Csibra, & Johnson, 2008) suggest that referential expectation is primarily elicited by ostensive-communicative contexts in infants, we modified the priming paradigm to provide an optimal stimulus environment for 9-month-olds. First, we presented infants with live rather than pre-recorded speech, which was occasionally accompanied by non-verbal referential gestures, like pointing and gazing towards upcoming objects. Second, we used the words as primes and the objects as probes because a known word is more likely to activate the associated semantic representation in pre-verbal infants than a picture of an object (cf. Xu, 2007). Third, instead of flashing a picture on the monitor, the object appeared in a dynamic fashion from behind a rotating occluder on a computer screen. Thus, while live speech and interaction ensured optimal conditions for speech processing of the prime (Kuhl, 2007), video presentation of the target objects made possible the accurate control of stimulus variables for measuring ERPs. Finally, because young infants display a preference for maternal speech (e.g. Cooper & Aslin, 1989), we used the mother as the speaker in one of the conditions. As auditorily-presented words have rarely been used as primes for visual probes in ERP studies (but see Nigam, Hoffman, & Simons, 1992, in a different paradigm), we also tested whether, using this paradigm, we could find a reliable N400 effect in adults.

Method

Participants

Twelve Hungarian native speaking adults (6 females, average age 38 years and 8 months, range: 25 to 56 years) and 28 infants raised in Hungarian families participated in the study. Infants were assigned to two conditions: 14 to the Mother-speech condition (5 females, average age 277 days, range: 269 to 286 days) and 14 to the Experimenter-speech condition

(6 females, 278 days, range: 266 to 285 days). One additional adult was excluded because of poor EEG impedance, and 21 additional infants were excluded because fussiness (11), insufficient trials (6), extensive body movements (2), poor impedance (1), or maternal interference (1).

Stimuli

We selected 15 object labels that, according to parental reports, two thirds of Hungarian 1-year-old infants recognize. We used corresponding pictures of objects to create 15 animated video-clips showing an occluder dropping to reveal one of the objects (Figure 1; see on-line Supplementary Materials for further details about the stimuli).

Procedure

Adults sat 70 cm in front of a CRT monitor. In each trial, they heard a pre-recorded word (the name of one of the 15 objects) from a loudspeaker behind the monitor, while a dynamic fixation stimulus was presented on top of an occluder. The duration of the auditory stimulus was between 419 and 784 ms (average 559 ms), which was followed by an interstimulus interval varying between 600 to 800 ms, during which the fixation stimulus ceased to move. Then the fixation stimulus disappeared and the occluder started to move down, revealing an object behind. This phase lasted for 480 ms. The object was fully visible for 1000 ms before the occluder screen rose, hiding the object again. Participants were presented with 240 trials in 4 blocks. In half of the trials the object corresponded to the preceding auditory word; in the other half they were inconsistent. Congruous and incongruous trials were presented equiprobably in pseudo-random order.

Infants sat on a highchair in front of a computer monitor. The infant's mother and an experimenter (E1) sat on small chairs at the two sides of the infant, both wearing headphones. At the beginning of each trial, the auditory stimulus was presented either to the mother (Mother-speech condition) or to E1 (Experimenter-speech condition), who repeated the word live for the infant. Mothers were instructed to speak to the child as they would do in everyday life, and they were allowed to gesture toward the monitor on which the occluder was seen. We asked them to utter the word at the very end of the sentence if they wanted to use complete expressions. In the Experimenter-speech condition, E1 talked to the infant, attempting to reproduce the words, intonation, and style of a yoked mother from the Mother-speech condition. In this design, E1 matched each individual mother from the Mother-speech condition for an infant in the Experimenter-speech condition.

Once the word was uttered and the infant attended the monitor, a second experimenter started a clip revealing an object behind the occluder, just as in the procedure for adult participants. Because of live presentation, the inter-stimulus interval between the word prime and the start of the visual stimulus varied across trials, and averaged about 2155 ms. Trials were presented as long as the infants were attentive. The position of the speaker was counterbalanced across subjects in both conditions. The behaviour of the infants was video-recorded throughout the session for offline trial-by-trial editing and for additional behavioral scoring. (Further details of the procedure are reported in the on-line Supplementary Materials.)

EEG Recording and Analysis

EEG was recorded by 128-channel Geodesic Sensor Nets at 500 Hz sampling rate. The EEG was segmented into 1700 ms long epochs including 200 ms preceding occluder motion. We considered time 0 the frame when the object started to become visible from behind the occluder (160 ms after motion onset). EEG segments were averaged to separate ERPs for word-congruous and word-incongruous objects, baseline-corrected to the first 200 ms of the segments and re-referenced to average reference. Based on previous reports that found N400 over parietal sites (Kutas & Hillyard, 1980), we identified as regions of interest the electrodes between C3 and P3 and between C4 and P4, on the left and right side, respectively (Figure 2). On the basis of visual inspection of the grand averages and existing literature on N400 latencies, adults' and infants' ERPs were analyzed as mean amplitude between 350 and 500 ms, and 500 and 650 ms, respectively, after object appearance (cf. Friedrich & Friederici, 2004; 2005b for N400 time windows in adults and infants). (Further details of the EEG procedure are reported in the on-line Supplementary Materials.)

Results

An analysis of variance (ANOVA) on the data from adults, with congruency (congruous vs. incongruous object) and hemisphere (left vs. right) as within-subjects factors, revealed that incongruous objects elicited a more negative N400 compared to the congruous objects ($F(1,11) = 7.50, p = .02, \eta^2_p = .41$). Although the effect was bigger on the right side, the interaction with hemisphere was not significant. This result demonstrates that a semantic congruency effect on the N400 can be elicited when an auditory word prime precedes an object image.

On average, infants contributed 20.3 and 20.3 trials to the congruous and incongruous ERPs, respectively, in the Mother-speech condition, and 19.4 and 19.9 trials in the Experimenter-speech condition. An ANOVA on the infant data with congruency and hemisphere as within-subjects factors and condition (Mother-speech vs. Experimenter-speech) as a between-subjects factor revealed an interaction between congruency and condition ($F(1,26) = 5.02, p = .03, \eta^2_p = .16$). This was due to the fact that the N400 was more negative to the incongruous than to the congruous object in the Mother-speech condition ($F(1,13) = 5.45, p = .036, \eta^2_p = .30$), but not in the Experimenter-speech condition ($F(1,13) = 0.29, p = .598, \eta^2_p = .02$) as reflected by two-way follow-up ANOVAs with congruency and hemisphere as factors (Figure 2). A congruency by hemisphere interaction in the three-way ANOVA also indicated that the effect was more pronounced over the right hemisphere ($F(1,26) = 4.39, p = .05, \eta^2_p = .15$).

In order to account for the difference in N400 between the conditions, we took five behavioural measures from infants: the number of times the speaker repeated the object name within a trial, whether the infant was looking to the mother or to E1 during the speech and whether the infant looked to the mother or to E1 after seeing the object. Although we tried to match the mothers' behaviour in the Experimenter-speech condition, mothers uttered the prime word slightly, but significantly, more than E1 did (1.07 vs. 1.01, $F(1,26) = 13.67, p = .001, \eta^2_p = .345$). As a control, we re-ran the statistics on ERPs that included only trials in which the word was uttered once in the Mother-speech condition, and we obtained the

same pattern of results as before. We did not find a significant difference between conditions in the number of times infants looked to the mother or to E1. However, infants in the Mother-speech condition tended to look more towards the mother than those in the other condition ($F(1,26) = 3.74, p = .064, \eta^2_p = .126$). There was no effect of any factor on how many times infants looked at the mother or at E1 after having seen the object.

Discussion

Our results suggest that infants as young as 9 months have a rudimentary receptive vocabulary. This has been previously suspected (Swingley, 2008), but difficult to prove. In agreement with our finding, a recent eye-tracker study by Bergelson and Swingley (in press) has shown that 6- to 9-month-olds can follow their mother verbal instructions and direct their gaze to objects. Six-month-old infants are also able to match novel words with arbitrary visual referents in a few trials by taking advantage of prosodic cues (Friedrich & Friederici, 2011; Shukla, et al., 2011). However, the electrophysiological signs of semantic priming disappeared after twenty-four hours, suggesting strong limitations in memory processes (Friedrich & Friederici, 2011). Our study shows that word-to-object priming occurs in 9-month-olds with familiar words when there is no requirement to learn new ones. Our method does not allow us to tell which infant understood which word, but our participants in the Mother-speech condition, as a group, seem to have activated the object features associated with the highly familiar words their mother uttered and matched them with the image that appeared on the screen in front of them. In this sense, infants in our study ‘understood’ their mother’s speech.

What is the nature of this understanding? In particular, is it possible that the word knowledge that infants displayed in this experiment goes beyond the formation of merely associative links between auditory and visual information, and reflects truly referential and semantic understanding of nouns? The kind of neuronal activation that we demonstrated here is correlational in nature and does not support unambiguous conclusions about the underlying processes. Nevertheless, we argue that several aspects of our results suggest that 9-month-olds appreciate the referential nature of words. First, the N400 component is commonly interpreted to reflect semantic processing by exhibiting lower amplitude to stimuli semantically (rather than grammatically or associatively) primed by the context (Kutas & Federmeier, 2011). The differential N400 responses to congruent and incongruent object images thus indicate that the words activated neural processes in infants that are correlated with the extraction of meaning from stimuli. Second, we found these effects despite the relatively long delay (> 2 seconds on average) between the uttered word and the appearance of the object. Earlier studies, which used synchronous presentation (with the object being visible during the presentation of the word), should have had a better chance to demonstrate associative audio-visual links, but failed to do so (Friedrich & Friederici, 2005b). Although pure associations could bridge temporal delays, they should work best with contiguous stimuli. In contrast, words can refer to absent referents, as they did in our study. Thus, finding differential brain activations for matches and mismatches between temporally separated stimuli supports the interpretation that a semantic, rather than an associative, link was formed between them.

Third, the success of this study was partly due to our effort to set up a situation in which infants had every reason to expect semantically interpretable referential expressions. Infants pay special attention to ostensive-communicative signals such as eye contact (Parise, Reid, Stets, & Striano, 2008), their own name (Parise, Friederici, & Striano, 2010), deictic reference, such as gaze (Hoehl, Wiese, & Striano, 2008), and pointing (Gredeback & Melinder, 2010) from early on. Ostensive communication has been suggested to generate referential expectation (Csibra, 2010) and expectation of co-reference for concurrent referential signals (Gliga & Csibra, 2009). While it is not clear why ostensive-referential signals would facilitate associative processes, they could support the inference that the words infants hear are semantically related to the object that would appear at the location referred by the speaker. Thus, our paradigm, which resembles closely the natural ‘joint attention’ interaction in which infants at this age are regularly engaged with their parents (Bakeman & Adamson, 1984), provided an optimal environment for measuring the effect of semantic priming in young infants.

Our results suggest that adopting the mother as the speaker also contributed to this optimal environment. Newborns prefer to listen to their mother’s voice compared to a stranger’s voice (DeCasper & Fifer, 1980), and recent ERP research confirmed that at the age of 4 months, infants respond faster and allocate more attention to the mother’s voice than an unfamiliar voice (Purhonen, Kilpeläinen-Lees, Valkonen-Korhonen, Karhu, & Lehtonen, 2005). In addition, the mother’s voice elicits more activation in language-relevant brain areas in newborns than a stranger’s (Beauchemin, et al., 2010). Our paradigm does not allow us to pinpoint the exact factors that made infants more responsive to the mother’s than to the experimenter’s communication. It could be that the familiar voice, intonation, or familiar verbal expression (e.g., ‘Look, a duck!’ vs. ‘Here comes the duck!’) helped them to recognize the situation as a naming game. It is also possible that 9-month-old infants had difficulties in recognizing the target word in the slightly different phonetic production by the experimenter. Because the speaker sat next to the infants and did not make always eye contact with them during her speech, they might not have recognized from the experimenter’s intonation alone that they were being addressed, while the mother’s voice alone could have achieved this effect. Further studies will test at what age or under what conditions infants detect semantic violations in a stranger’s speech. Nevertheless, the functional nature of the N400, and the similarity of the effects we found in adults and infants, suggests semantic priming by, and referential understanding of, familiar words at 9 months of age. This finding supports the view that the referential nature of speech may not have to be learned by human infants, but it may be expected and exploited by them during language acquisition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by an ERC Advanced Investigator Grant (OSTREFCOM) to G. Csibra. We thank B. Kollod, M. Toth, and A. Volein for assistance, and H. Bortfeld, M. Chen, M. Friederich, A. D. Friederici, T. Gliga, A. M. Kovacs and O. Mascaro for discussions.

References

- Bakeman R, Adamson LB. Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*. 1984; 55(4):1278–1289. [PubMed: 6488956]
- Beauchemin M, González-Frankenberger B, Tremblay J, Vannasing P, Martínez-Montes E, Belin P, et al. Mother and Stranger: An Electrophysiological Study of Voice Processing in Newborns. *Cerebral Cortex* (New York, NY: 1991). 2010
- Bergelson E, Swingle D. At six to nine months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*. (in press).
- Cooper RP, Aslin RN. The language environment of the young infant: implications for early perceptual development. *Canadian Journal of Psychology*. 1989; 43(2):247–265. [PubMed: 2486498]
- Csibra G. Recognizing Communicative Intentions in Infancy. *Mind & Language*. 2010; 25(2):141–168.
- DeCasper AJ, Fifer WP. Of human bonding: newborns prefer their mothers' voices. *Science* (New York, NY). 1980; 208(4448):1174–1176. [PubMed: 7375928]
- Friedrich M, Friederici AD. N400-like semantic incongruity effect in 19-month-olds: processing known words in picture contexts. *Journal of Cognitive Neuroscience*. 2004; 16(8):1465–1477. [PubMed: 15509391]
- Friedrich M, Friederici AD. Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*. 2005a; 16(6):653–656. [PubMed: 15812327]
- Friedrich M, Friederici AD. Phonotactic knowledge and lexical-semantic processing in one-year-olds: brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*. 2005b; 17(11):1785–1802. [PubMed: 16269114]
- Friedrich M, Friederici AD. Maturing brain mechanisms and developing behavioral language skills. *Brain and Language*. 2010; 114(2):66–71. [PubMed: 19665783]
- Friedrich M, Friederici AD. Word Learning in 6-Month-Olds: Fast Encoding-Weak Retention. *Journal of Cognitive Neuroscience*. 2011
- Gleitman LR, Gleitman H. A Picture Is Worth a Thousand Words, but That's the Problem: The Role of Syntax in Vocabulary Acquisition. *Current Directions in Psychological Science*. 1992; 1(1):31–35.
- Gliga T, Csibra G. One-year-old infants appreciate the referential nature of deictic gestures and words. *Psychological Science*. 2009; 20(3):347–353. [PubMed: 19207689]
- Gredebäck G, Melinder A. The development and neural basis of pointing comprehension. *Social Neuroscience*. 2010; 5(5-6):441–450. [PubMed: 20162491]
- Hoehl S, Wiese L, Striano T. Young infants' neural processing of objects is affected by eye gaze direction and emotional expression. *PLoS ONE*. 2008; 3(6):e2389. [PubMed: 18545689]
- Hollich GJ, Hirsh-Pasek K, Golinkoff RM, Brand RJ, Brown E, Chung HL, et al. Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*. 2000; 65(3):i–vi. 1–123. [PubMed: 12467096]
- Kuhl PK. Early language acquisition: cracking the speech code. *Nature Review Neuroscience*. 2004; 5(11):831–843. [PubMed: 15496861]
- Kuhl PK. Is speech learning 'gated' by the social brain? *Developmental Science*. 2007; 10(1):110–120. [PubMed: 17181708]
- Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*. 2011; 62:621–647.
- Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*. 1980; 207(4427):203–205. [PubMed: 7350657]
- Nigam A, Hoffman J, Simons R. N400 to Semantically Anomalous Pictures and Words. *Journal of Cognitive Neuroscience*. 1992; 4(1):15–22. [PubMed: 23967854]
- Parise E, Friederici AD, Striano T. “Did you call me?” 5-month-old infants own name guides their attention. *PLoS One*. 2010; 5(12):e14208. [PubMed: 21151971]

- Parise E, Reid VM, Stets M, Striano T. Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience*. 2008; 3(2):141–150. [PubMed: 18633855]
- Purhonen M, Kilpeläinen-Lees R, Valkonen-Korhonen M, Karhu J, Lehtonen J. Four-month-old infants process own mother's voice faster than unfamiliar voices--electrical signs of sensitization in infant brain. *Brain Research Cognitive Brain Research*. 2005; 24(3):627–633. [PubMed: 16099371]
- Robinson, CW.; Howard, EM.; Sloutsky, VM. The nature of early word comprehension: Symbols or associations?. In: Bara, BG.; Barsalou, L.; Bucciarelli, M., editors. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*; Mahwah, NJ: Erlbaum; 2005. p. 1883-1888.
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science*. 1996; 274(5294):1926–1928. [PubMed: 8943209]
- Senju A, Csibra G. Gaze following in human infants depends on communicative signals. *Current Biology*. 2008; 18(9):668–671. [PubMed: 18439827]
- Senju A, Csibra G, Johnson MH. Understanding the referential nature of looking: infants' preference for object-directed gaze. *Cognition*. 2008; 108(2):303–319. [PubMed: 18371943]
- Shukla M, White KS, Aslin RN. Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(15):6038–6043. [PubMed: 21444800]
- Swingley D. The roots of the early vocabulary in infants' learning from speech. *Current Directions in Psychological Science*. 2008; 17(5):308–311. [PubMed: 20523916]
- Waxman SR, Gelman SA. Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*. 2009; 13(6):258–263. [PubMed: 19447670]
- Xu F. Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*. 2007; 11(9): 400–406. [PubMed: 17698404]

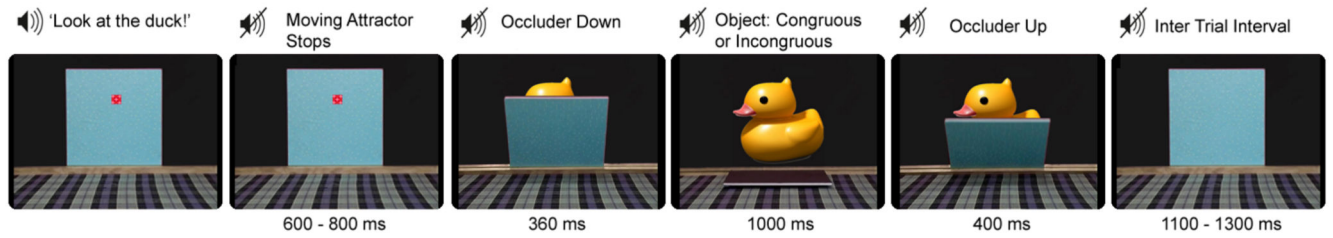


Figure 1.

An example of the stimuli in a trial for infants. The word is uttered live by either the mother or an experimenter. ERPs were time-locked to the first video frame revealing the top of the object behind the occluder.

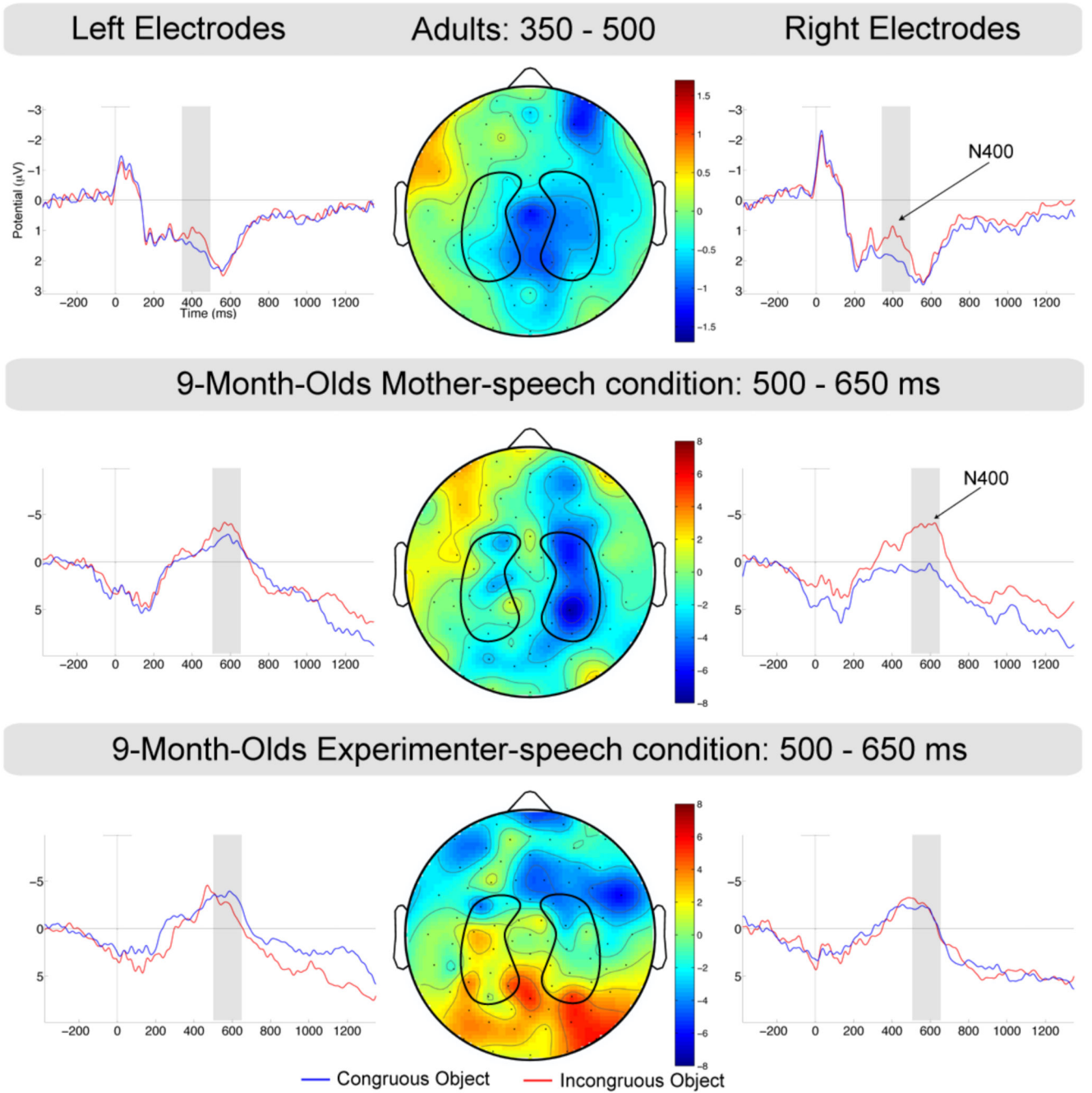


Figure 2. ERPs over left and right regions of interest (marked by black contours on the scalp maps). The scalp maps depict the spatial distribution of the difference between word-incongruous and word-congruous objects in the given time windows for adults and infants.