Fourteen-Month-Old Infants Track the Language Comprehension of Communicative Partners

Running title: Infants process others' linguistic understanding

Authors: Bálint Forgács[a,b,c,1], Eugenio Parise[d,e], Gergely Csibra[e,f], György Gergely[e], Lisa Jacquey[a,b], Judit Gervain[a,b]

Author affiliations:

[a] Laboratoire Psychologie de la Perception (LPP), Université Paris Descartes, 45, rue des Saints-Pères, 75006 Paris, France

[b] Laboratoire Psychologie de la Perception (LPP), CNRS, 45, rue des Saints-Pères, 75006 Paris, France

[c] Department of Cognitive Psychology, Eötvös Loránd University (ELTE), Izabella utca 46, 1064 Budapest, Hungary

[d] Department of Psychology, Lancaster University, Lancaster LA1 4YF, United Kingdom

[e] Cognitive Development Center (CDC), Department of Cognitive Science, Central European University (CEU), Nádor utca 9, 1051 Budapest, Hungary

[f] Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London WC1E 7HX, United Kingdom

[1]*Corresponding author:* Bálint Forgács: forgacs.balint@ppk.elte.hu

technical help.

Research Highlights

- 14-month-olds show an N400 effect even when only an observer encounters semantic incongruity (an object mislabeled), themselves not.

- Infants exhibit similar neural markers when following the linguistic understanding of a social partner as they show when processing word meaning.

- Results suggest that infant Theory-of-Mind goes beyond the attribution of false beliefs about object location as it allows attribution of miscomprehension as well.

Abstract

Infants employ sophisticated mechanisms to acquire their first language, including some that rely on taking the perspective of adults as speakers or listeners. When do infants first show awareness of what other people understand? We tested 14-month-old infants in two experiments measuring event-related potentials. In Experiment 1, we established that infants produce the N400 effect, a brain signature of semantic violations, in a live object naming paradigm in the presence of an adult observer. In Experiment 2, we induced false beliefs about the labelled objects in the adult observer to test whether infants keep track of the other person's comprehension. The results revealed that infants reacted to the semantic incongruity heard by the other as if they encountered it themselves: they exhibited an N400-like response, even though labels were congruous from their perspective. This finding demonstrates that infants track the linguistic understanding of social partners.

Introduction

In order to be able to appraise the communicative value of words and to recognize their meaning, it is important to assume that others understand them in the context of their beliefs. Language is not the only means for information transmission, but it is a particularly efficient way to change the mental states of social partners. If human communication is fundamentally inferential in nature (Sperber & Wilson, 1995), communicators should ordinarily rely on their Theory-of-Mind (ToM) in monitoring and guiding their interlocutors' comprehension of linguistic utterances.

Some evidence suggests that, when evaluating others' interpretation of utterances, people take into account the information available to their partners. Keeping the common ground of a conversation has been proposed to drive acknowledgements, turn taking, and continued attention (e.g., Clark & Brennan, 1991), while people also seem to be influenced by the visual perspective of a conversational partner in communicational situations (e.g., Keysar, Barr, Balin, & Brauner, 2000). Rueschemeyer, Gardner, and Stoner (2015) devised a "joint comprehension task", during which they recorded the event-related potentials (ERPs) of participants, who read sentences with congruent or incongruent endings on a computer screen together with a confederate. The N400 ERP component is a well-established neurophysiological indicator of semantic comprehension, which is sensitive to semantic violations, such as incongruent ending of sentences (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980). Intriguingly, Rueschemeyer and colleagues (2015) found that when adult participants' – or even adolescents' (Westley, Kohút, & Rueschemeyer, 2017) – task was to monitor the comprehension of the confederate, they produced an increased N400 for sentence endings that were incongruent for the confederate only. Crucially, only the participant, but not the confederate, had access to additional contextual information played back on headphones. This finding indicates that judging others' semantic evaluation of linguistic utterances relies on tracking, and taking into account, the interlocutors' beliefs.

Infants seem to be particularly well equipped for learning about linguistic and nonverbal communication. Newborns are already sensitive to the abstract structure (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008) and certain acoustic characteristics (Gervain & Mehler, 2010) of language. They exhibit specific sensitivity to ostensive signals (Csibra, 2010), such as eye-gaze, infant-directed speech, or being addressed by their name, as measured by electroencephalography (EEG) around 5 months of age (Parise & Csibra, 2013; Parise, Friederici, & Striano, 2010), and by functional near-infrared spectroscopy (fNIRS) around 6 months of age (Grossmann, Parise, & Friederici, 2010; Lloyd-Fox, Széplaki-Köllőd,

Yin, & Csibra, 2015). By 6 months, infants expect speech to transmit information between agents even about invisible entities, such as intentions or preferences (Vouloumanos, Martin, & Onishi, 2014), and they already possess a small receptive vocabulary (Bergelson & Aslin, 2017; Bergelson & Swingley, 2012). Thus, the communicative value of speech seems to be evident for infants well before they produce their first words.

The capacity to represent others' beliefs, crucial to be able to engage in inferential communication (Sperber & Wilson, 1995), seems to be present during the first year of life (Baillargeon, Scott, & He, 2010), perhaps even as early as 6 or 7 months of age (Kovács, Téglás, & Endress, 2010; Southgate & Vernetti, 2014). Therefore, basic elements of ToM, in particular the ability to track others' perspective and mental states, can contribute to language acquisition, especially to lexical development. Shortly after the first birthday, language comprehension seems to flourish indeed: word learning becomes more efficient and flexible (Werker, Cohen, Lloyd, Casasola, & Stager, 1998), words map onto concepts (Yin & Csibra, 2015), and the vocabulary is extended to abstract words (Bergelson & Swingley, 2013). The N400 can also be elicited around this age, but initially only under specific conditions. For example, high word producer 12-month-olds show it, but low word producers do not (Friedrich & Friederici, 2010). Infants as young as 9-month-olds exhibit an N400 if they hear objects named by their own mother, but not by an experimenter (Parise & Csibra, 2012), or if they are exposed to extensive familiarization with word-object pairings (Junge, Cutler, & Hagoort, 2012). By 14 months of age the N400 can be evoked reliably without these additional aids (Friedrich & Friederici, 2005, 2008), suggesting that semantic comprehension abilities become less dependent on situational aspects. Finally, at the same age, there is recognition of at least some sort of "common ground" (Moll, Richter, Carpenter, & Tomasello, 2008) in social interactions. These developmental steps seem to be critical for effective inferential communication, allowing infants to start to take into account what others mean and understand.

The experiments reported here addressed the question whether 14-month-old infants take the perspective of a communicative partner when they evaluate linguistic utterances, in particular, labels applied to familiar objects. Since infants show evidence of visual perspective taking from at least 8 months of age (Kampis, Parise, Csibra, & Kovács, 2015), track others' beliefs about objects, and evaluate referential expressions semantically, they should be able to judge how a communicative partner with a different perspective would interpret object labels.

Experiment 1

This experiment aimed to establish whether infants would display an N400 effect in a live setting where another person is also present. Previous studies with infants have shown semantic incongruity effects using video presentations, but never in live communicative contexts. During a presentation of familiar objects in a puppet theater, we recorded the EEG of 14-month-old infants, while a hand pointed to and a voice named the objects in the presence of an adult Observer. We measured infants' ERPs time-locked to the onset of the word label, which was either congruent or incongruent with the object for both parties.

Method

  Participants. Eighteen full-term, 14-month-old infants (mean age: 444 days; range: 428-457 days) participated in this experiment. They were all French monolinguals (hearing no more than 15% of any other languages according to parental report). The sample size – for both experiments – was determined based on previous experiments of the authors and studies investigating infant N400 effects (Friedrich & Friederici, 2005, 2008, 2010; Parise & Csibra, 2012; Junge et al., 2012). An additional 68 infants were excluded from the final data analysis (21% inclusion rate). The main reason for exclusion was not meeting the criteria of having at least 10 artifact free trials per condition (Stahl, Parise, Hoehl, & Striano, 2010) (n=30), primarily because of eye-movements (e.g., switching gaze from the Observer to the object or the other way around during the measurement period), blinking, or body movements. Further reasons for exclusion were fussiness (n=25), refusing the EEG cap (n=3), experimenter error (n=2), or providing too noisy ERPs (n=8)[1]. Such an attrition rate is not unprecedented in the infant EEG literature (cf. Parise, Reid, Stets, & Striano, 2008). It can be attributed to natural infant behavior, such as visual exploration of the scene, and also to the length of the experiment: a single trial lasted about 20-25 seconds because of the elaborate live experimental procedure, and together with the necessary short breaks, an infant needed to stay attentive for about 25-35 minutes to produce a sufficient number of artifact free trials. We believe that participant exclusion was unbiased with respect to any stable individual differences that would be relevant to the capabilities studied here. An additional 4 infants were tested during the training of the assistants playing the role of the Observer; the first

---

[1] Individual ERPs were considered too noisy if they did not conform to a few classical indicators of successful recordings, such as: a flat pre-stimulus baseline, initially aligned ERPs across conditions, and a mild positive deflection frontally and a mild negative deflection parietally, characteristic of auditory processing.

couple of experiments with each new assistant were considered practice sessions, the data of which were not be included.

Materials. We used the audio recording of 15 object labels in French, and selected corresponding real-world objects that we suspected that infants knew the labels for (e.g., bunny, cup, shoe, duck, etc.). These word-object pairs were adapted from Parise and Csibra (2012), and were confirmed via the French version of the CDI (Kern, 2007). Additionally, we recorded seven short utterances to get infants' attention, such as "Look!", "How interesting!", "Listen!" etc., in French, which were played back when needed (see Apparatus & Procedure sections below). A female speaker's voice was recorded for all audio materials, she uttered the words and the sentences in a language-appropriate infant-directed intonation.

Apparatus. Infants sat on their parent's lap in front of a small puppet theater. A research assistant, playing the role of the Observer, sat facing them on the other side of the theater stage, behind a 30 cm tall mobile occluder halfway on the stage, and a curtain that she could open and close. A second experimenter sat on the left side of the stage behind a curtain, hidden to the infant, manipulated the occluder, and placed and replaced objects according to instructions displayed on a screen not visible to the participants. We used E-prime 2.08 software (Psychology Software Tools, Pittsburgh, PA) to control sound playback and to provide EEG triggers at the onset of word labels. We also video-recorded the infants using a camera to check during later analysis for eye-blinks, head and body movements, and to monitor their attention trial by trial.

Procedure. Before the experiment commenced, we explained the procedure to the parents, and obtained their informed consent. Ethics approval was obtained from the Ethics Committee of Université Paris Descartes. Parents were instructed to look downwards to the floor throughout the experiment and not to interact with or talk to their children, unless they felt it was necessary to soothe them. (For further details see SI Procedure.)

In each trial (see Figure 1), the Observer first opened the curtains and established eye contact with the baby. An object (e.g., an apple) was placed in front of the occluder, so that the baby could see it while it was hidden from the Observer. Then the occluder was lowered so that the object was revealed to the Observer. She attempted to make eye-contact again with the infant by looking up from the object. This gaze alternation was repeated until the Observer was certain that the infant knew that she had seen the object. If the infant was not attentive, or fixated the object without looking at the Observer, she played pre-recorded attention getting utterances using a hidden button box and, if necessary, manipulated the object briefly. The occluder was then raised, and the Observer turned 90° to the left so that

her eyes were behind the curtain on the side, and the object was replaced with another object (e.g., a car) by the second experimenter from behind the curtain. When the second object was in place, the Observer turned back, and the second object was revealed to the Observer by lowering the occluder. (We included this change-of-object procedure in order to match the stimuli closely to the one to be used in Experiment 2.) The Observer again made sure that the infant noticed that she attended to the object. The second experimenter then raised the occluder again, pointed at the object from behind the curtain, and, using a hidden button box, the Observer played a recorded label that was, depending on the condition, either congruent or incongruent with the object. We took care to present the audio stimulus only when the infant was engaged. For two seconds following the playback, the Observer looked at the infant neutrally, while the pointing hand stood still, after which the object was removed, and the Observer closed the curtain. A two-second break indicated the end of the trial. The next trial started by opening the curtain. The order of objects and labels across trials were determined by a pseudo-random sequence, in which the same object did not appear more than once in three consecutive trials, and there were no more than two congruent or incongruent trials in a row. A video recording of the procedure is provided in the SI Video.

EEG recording and analysis. We recorded the EEG continuously via a 128-channel Geodesic Sensor Net at 500 Hz sampling rate. Raw data was filtered with a 0.3 Hz high-pass, and 30 Hz low-pass filter, and segmented into 1200 ms epochs starting at 200 ms before the onset of the audio playback in each trial. Artifact detection algorithms (e.g., blinks, eye-movements, bad channels, etc.) were used to exclude bad segments automatically, and video recordings confirmed that the infants attended the scenes in the retained segments. Infants must have had at least 10 artifact-free segments per condition to be included in the final data analysis (Stahl et al., 2010). Bad channels were replaced by spherical spline interpolation, segments were averaged separately for the congruent and incongruent conditions, baseline-corrected to the 200 ms preceding word onset, and re-referenced to the average reference. Raw EEG data are available upon request. When relevant, Mauchly's test was used to test for sphericity violations, and whenever necessary the Greenhouse-Geisser correction was applied. Statistical reports and effect size calculations follow the recommendations of Lakens (2013). (Further details are reported in SI EEG Recordings.)

Results

Based on previous findings (Friedrich & Friederici, 2005, 2008, 2010), we identified a time window between 400 and 600 ms, and 13 parietal electrodes as the region of interest (ROI)

for confirmatory statistical analyses. To quantify N400 activation, we computed the average amplitude within this time window and ROI. We found that the N400 was more negative in the incongruent ($M = -13.8$ μV, $SD = 6.88$) than in the congruent ($M = -10$ μV, $SD = 6.86$) condition, $t(17) = -2.14$, $p = .047$, 95% CI [-0.6 μV, -7.51 μV], Hedges's $g_{av} = .53$ (Figure 2). This confirmed that infants detected when labels did not match the object on the scene. Grand-average topographical ERP plots for all electrode sites can be viewed in Figure S1.

Discussion

Our results demonstrate that 14-month-olds display signatures of semantic comprehension in a live setting, when attending object naming together with an adult. We found an N400 effect with a timing and distribution very similar to a typical infant N400.

Experiment 2

This study addressed our main question: whether infants take into consideration others' mental states to follow their linguistic comprehension. We investigated whether 14-month-olds track word-referent relations as understood by their communicative partners, and whether the underlying computations rely on the semantic content of object labels. We expected that an N400 effect would reflect infants' evaluation of the semantic processing of object-label pairs from an observer's point of view.

In this experiment, the object name uttered was always congruent with the object from the perspective of the infant, but in half of the trials it was incongruent from the perspective of the Observer. We achieved this by inducing a false belief in the mind of the Observer about the identity of the object in the Observer-incongruent condition (Figure 1). The Observer-congruent condition was identical to the one of the first experiment, that is, the naming was congruent from both the Observer's and the infant's perspective.

Method

The methods of Experiment 2 closely matched to those of Experiment 1 with the following differences.

Participants. Eighteen full-term, French monolingual 14-month-olds (mean age: 443 days; range: 425-456 days) participated in the experiment. An additional 62 infants were excluded from the final data analysis (23% inclusion rate). The main reasons for exclusion were: not meeting the criteria of having at least 10 clean trials per condition (n=27), fussiness

(n=17), refusing the EEG cap (n=2), experimenter error (n=4), or providing too noisy individual grand averages (n=12). Further 8 infants participated in the training sessions of the assistants.

Procedure. In this experiment, the label was always congruent with the object on the scene, and the two conditions diverged from the point when the Observer turned back following the object change. In Observer-congruent trials, the second object was revealed to the Observer by lowering the occluder, just as in Experiment 1. When the object was revealed to the Observer, both the participant and the Observer had a true belief about the identity of the object, and the label was congruent from both of their perspectives. In contrast, in Observer-incongruent trials, the occluder was *not* lowered after the Observer turned back following the object change. The second experimenter pointed at the object from behind the curtain, and the Observer played back the recorded label of the second object. When the object was not revealed to the Observer, the naming was congruent with the object only for the infant, but not for the Observer, because she must have had no idea about the identity of the second object, and falsely believed that the first object was labeled, from her perspective, incongruently.

Results

The amplitude of the N400 was analyzed the same way as in Experiment 1. We found significantly greater negativity to the Observer-incongruent ($M = -13$ μV, $SD = 7.51$) as opposed to the Observer-congruent ($M = -8.42$ μV, $SD = 6.84$) condition, $t(17) = -2.46$, $p = .025$, 95% CI [-0.65 μV, -8.55 μV], Hedges's $g_{av} = .61$ (Figure 2), indicating that the infants detected the incongruency for the Observer even if the labels they heard and the objects they saw were congruent with each other. To compare the findings of the two experiments we entered the data in a 2x2 ANOVA with Condition (congruent, incongruent) as within-subject factor, and Experiment (1 & 2) as between-subject factor. The main effect of Condition was significant $F(1, 34) = 10.6$, $p = .003$, $\eta_p^2 = .24$, but neither the main effect of Experiment, nor the interaction between Condition and Experiment was (all $ps > .55$).

Based on the visual inspection of the waveforms, we also identified a sustained frontal negativity starting from around 600-700 ms after stimulus onset (Figure 3). We ran exploratory statistical analyses over a frontal ROI consisting of 13 electrodes. We compared the mean amplitudes of 100 ms time windows between 600-1000 ms after stimulus onset for the two conditions in a 2x4 ANOVA with Condition and Time Window as within-subject

factors. We found a significant main effect of Condition $F(1, 17) = 9.75$, $p = .006$, $\eta_p^2 = .36$ and of Time Window $F(1, 51) = 15.5$, $p < .001$, $\eta_p^2 = .48$, and a significant interaction of the two $F(2, 34) = 7.34$, $p = .002$, $\eta_p^2 = .3$. Planned comparisons revealed no difference between 600-700 ms, $t(17) = -1.89$, $p = .076$, Hedges's $g_{av} = .49$, but significant differences between 700-800 ms, $t(17) = -2.71$, $p = .015$, Hedges's $g_{av} = .71$; 800-900 ms, $t(17) = -3.52$, $p = .003$, 95% Hedges's $g_{av} = 1.02$; and 900-1000 ms, $t(17) = -3.51$, $p = .003$, Hedges's $g_{av} = .91$ (Figure 3). Grand-average topographical ERP plots for all electrode sites can be found in Figure S2, and further analyses in SI Results.

Discussion

Our results demonstrate that 14-month-olds spontaneously evaluate the verbal understanding of communicative partners. The difference found between Observer-congruent and Observer-incongruent conditions indicates that infants track the mental states of social partners, keep such attributed representations updated, and use them to assess others' semantic processing of words in relation to referents. Intriguingly, the timing and distribution of the N400 effect was very similar to what we found in Experiment 1, although the infants themselves should not have experienced any semantic incongruities between labels and objects. Such an incongruity must have originated from the semantic evaluation of the label-object match from the Observer's perspective, who must have believed that a different object was being labelled in Observer-incongruent trials. Thus, our study provided indirect evidence for attributing false beliefs to the Observer.

The frontal effect, which we did not predict but found in Experiment 2, is comparable to previous findings regarding ToM by Liu and colleagues (2009), although it is almost a second earlier than for preschoolers, and is in the time window of adult mentalizing processes (Liu, Sabbagh, Gehring, & Wellman, 2004, 2009). In contrast to typical false belief tasks, the potentially communicative nature of our design, and specifically the live setting we adopted, could have induced online monitoring of beliefs of real partners, which could underlie infants' adult-like speed of mentalizing. Our results also speak to the debate whether there are two ToM systems: an ancient yet efficient one available first, and a flexible one developing later (Low, Apperly, Butterfill, & Rakoczy, 2016); or a single system, fully capable of attributing mental states early on (Baillargeon et al., 2010; Buttelmann, Suhrke, & Buttelmann, 2015; Kovács, 2016). Our finding, in particular the early frontal effect, suggests that similar systems could be at work in 14-month-olds as in adults. Moreover, implicit false belief tasks are often portrayed as capable of demonstrating false belief attribution about

location, but not about identity, and it has been argued that only the latter would be evidence for the attribution of truly representational mental states (Low et al., 2016). Our results indicate that 14-month-old infants can attribute false beliefs about something that is not identity, but is not location either: the existence of an object belonging to a specific basic level category at a certain location.

General Discussion

The N400 ERP component is thought to reflect a violation of semantic evaluation (Kutas & Federmeier, 2011; Brouwer, Fitz, & Hoeks, 2012). Experiment 1 demonstrated that such an effect can be recorded from infants in a live setting when they match linguistic labels to objects. However, we also found this effect in Experiment 2, where no semantic violation was apparent from the perspective of the infants – as if the Observer's knowledge state were a constitutive part of the semantic context. Our finding therefore goes beyond the "social N400" effect reported by Rueschemeyer and colleagues (2015) and Westley and colleagues (2017): the explicit nature of their task could have invited top down control mechanisms to monitor semantic information available to the partner. In contrast, our task was not only implicit, as it was not possible to give instructions to infants, but the information asymmetry arose from differing perspectives, not from differential availability of semantic information. In other words, the infants did not need to assess the linguistic content of their social partner's belief, nevertheless, they spontaneously did so. Given the relatively high attrition rate, further research may be needed to replicate these results.

One potential source of explanation of this finding is that infants might rely on the same neural and cognitive mechanisms to accomplish a task for themselves as they do for others when processing mental states (Kampis et al., 2015). Consequently, infants could be utilizing comprehension mechanisms devoted to detecting semantic incongruities when they are motivated to guess how another person could have understood an utterance. A further potential explanation of our finding is that the N400 effect reflects a violation relative to the general, shared semantic context, a sort of common ground (Moll et al., 2008) or cognitive environment (Sperber & Wilson, 1995). Finally, it is also possible that, when they evaluate semantic relations, infants give priority to the communicators' mental states (Southgate, 2014). The current experiment cannot sufficiently adjudicate among these accounts, but all these explanations share the requirement of tracking the semantic information available to communicative partners.

In sum, we found a neural marker of semantic incongruity detection in infants when

their social partners, but not themselves heard objects labeled incorrectly. We also found a later frontal effect, similarly to studies investigating ToM, with a timing comparable to that of adults. Taken together, infants appear to be not only well prepared to learn language and track the mental states of social partners, but also to appreciate language at its full potential in terms of inferential communication: as a system where words refer to things, have culturally fixed senses, and enable us to change each other's mental representations. Even before they start to speak, infants monitor the (mis)comprehension of communicative partners, raising the possibility that infants have a much more social take on language in general than previously thought.

References

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110-118.

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences, 114*(49), 12916-12921.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391-397.

Berger, A., Tzur, G., & Posner, M. I. (2006). Infant brains detect arithmetic errors. *Proceedings of the National Academy of Sciences*, *103*(33), 12649-12653.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127-143.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology*, *131*, 94-103.

Csibra G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25*(2), 141-168.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on socially shared cognition, 13*, 127-149.

Friedrich, M., & Friederici, A. D. (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*, *16*(6), 653-656.

Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757-1761.

Friedrich, M., & Friederici, A. D. (2010). Maturing brain mechanisms and developing behavioral language skills. *Brain and Language*, *114*(2), 66-71.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222-14227.

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, *61*, 191-218.

Grossmann, T., Parise, E., & Friederici, A. D. (2010). The detection of communicative signals directed at the self in infant prefrontal cortex. *Frontiers in Human Neuroscience*, *4*.

Junge, C., Cutler, A., & Hagoort, P. (2012). Electrophysiological evidence of early word learning. *Neuropsychologia*, *50*(14), 3702-3712.

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. In *Proc. R. Soc. B, 282*, 20151683.

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227-250.

Kovács, Á. M. (2016). Belief files in theory of mind reasoning. *Review of Philosophy and Psychology*, *7*(2), 509-527.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830-1834.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology, 62*, 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11*(1), 32-38.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology, 4*, 863.

Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2004). Decoupling beliefs from

reality in the brain: an ERP study of theory of mind. *NeuroReport*, *15*(6), 991-995.

Liu, D., Sabbagh, M. A., Gehring, W. J., & Wellman, H. M. (2009). Neural correlates of children's theory of mind development. *Child Development*, *80*(2), 318-326.

Lloyd-Fox, S., Széplaki-Köllőd, B., Yin, J., & Csibra, G. (2015). Are you talking to me? Neural activations in 6-month-old infants in response to being addressed during natural interactions. *Cortex*, *70*, 35-48.

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: a primer on the two-systems account. *Child Development Perspectives*, *10*(3), 184-189.

Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-Month-Olds Know What "We" Have Shared in a Special Way. *Infancy*, *13*(1), 90-101.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science*, *23*(7), 728-733.

Parise, E., & Csibra, G. (2013). Neural responses to multimodal ostensive signals in 5-month-old infants. *PloS One*, *8*(8), e72360.

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.

Parise, E., Reid, V. M., Stets, M., & Striano, T. (2008). Direct eye contact influences the neural processing of objects in 5-month-old infants. *Social Neuroscience*, *3*(2), 141-150.

Rueschemeyer, S. A., Gardner, T., & Stoner, C. (2015). The Social N400 effect: how the presence of other listeners affects language comprehension. *Psychonomic Bulletin & Review*, *22*(1), 128-134.

Southgate, V. (2014). Early manifestations of mind reading. In S. Baron-Cohen, H. Tager-Flusberg & M. Lombardo (Eds.), Understanding Other Minds, 3rd Edition. Oxford University Press

Southgate, V., & Vernetti, A. (2014). Belief-based action prediction in preverbal infants. *Cognition*, *130*(1), 1-10.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition.* Oxford: Blackwell (second edition with Postface).

Stahl, D., Parise, E., Hoehl, S., & Striano, T. (2010). Eye contact and emotional face processing in 6-month-old infants: Advanced statistical methods applied to event-related potentials. *Brain and Development*, *32*(4), 305-317.

Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, *17*(6), 872-879.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, *34*(6), 1289.

Westley, A., Kohút, Z., & Rueschemeyer, S. A. (2017). "I know something you don't know": Discourse and social context effects on the N400 in adolescents. *Journal of experimental child psychology, 164*, 45-54.

Yin, J., & Csibra, G. (2015). Concept-based word learning in human infants. *Psychological Science*, *26*(8), 1316-1324.

Figure Legends

Figure 1. Experimental procedure. These events were performed in a live setting while the Observer was interacting with infant participants.

Figure 2. The N400 effect in Experiment 1 & Experiment 2. The upper panels show grand-average waveforms over the parietal ROI. Negative is plotted up; time 0 is the onset of the audio playback. Green shadings indicate the time window of the infant N400 (400-600 ms). Topographical maps below show scalp distributions of ERP differences in the N400 time window. Colder colors indicate greater negativities. Red rings illustrate the ROI over which the N400 amplitude was quantified.

Figure 3. Frontal ERP effects in Experiment 1 & Experiment 2. Upper panels show grand-average waveforms over the frontal ROI. Negative is plotted up; time 0 is the onset of the audio playback. Green shadings indicate where the frontal effect was significant (700-1000 ms). Topographical maps below show ERP differences in the 700-1000 ms time window for the two experiments. Colder colors indicate greater negativities. Red rings show the ROI over which the frontal negativity was quantified.

Experimental Procedure

Curtain opens, trial starts — 1-2 seconds
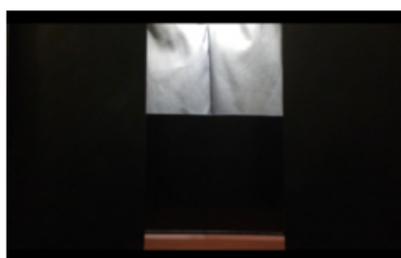
First object placed in front of occluder — 2-3 seconds
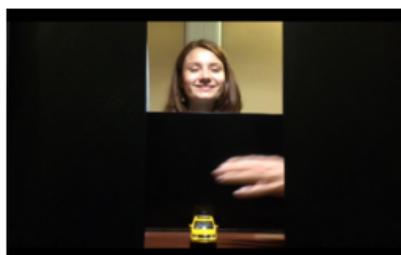
Occluder opens, observer notices first object — 2-3 seconds
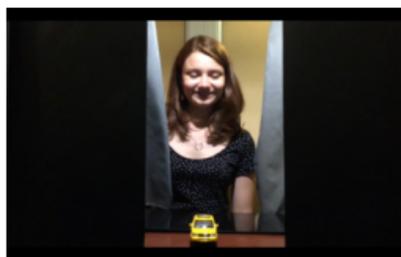
Observer turns away, first object replaced with a second one — 5-7 seconds
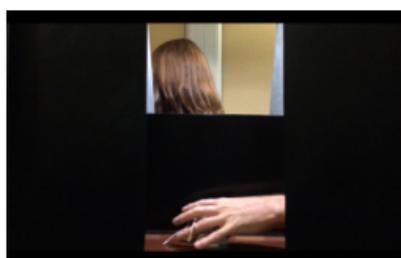
Observer turns back, eye-contact re-established — 1-2 seconds
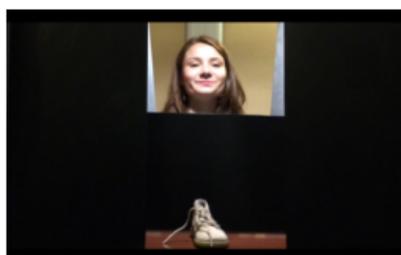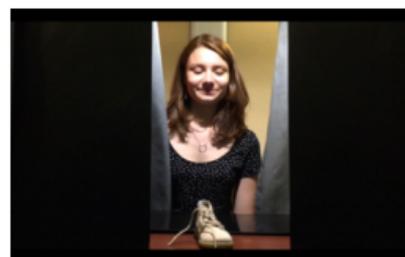
Experiment 1 and the Observer-congruent condition of Experiment 2
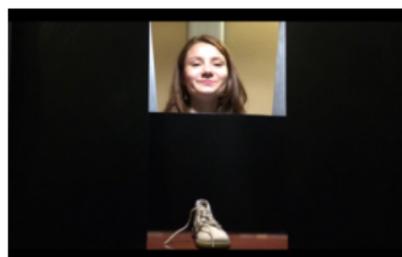
Observer-incongruent condition of Experiment 2
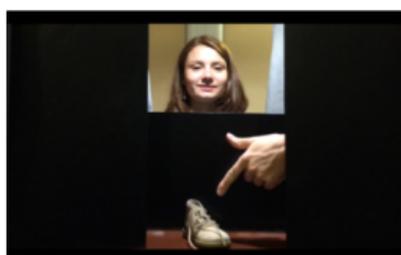
Pointing and audio playback of the object-label. The label is either congruent or incongruent with the object (Experiment 1), or is always congruent with it (Experiment 2).

Experiment 1
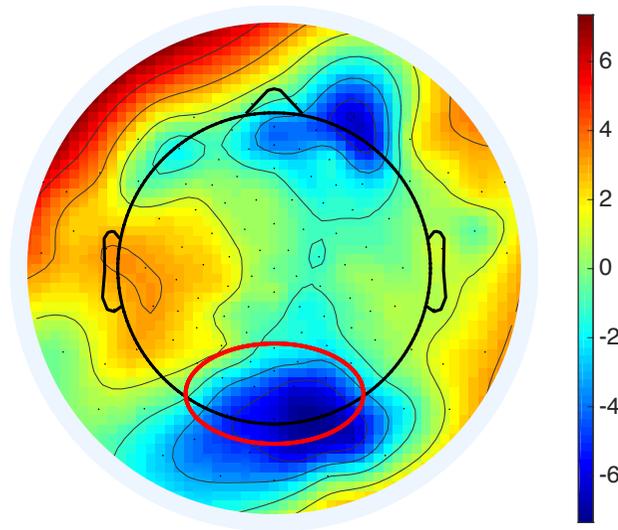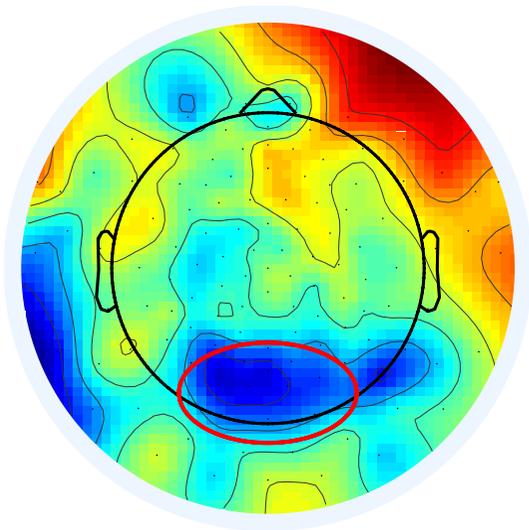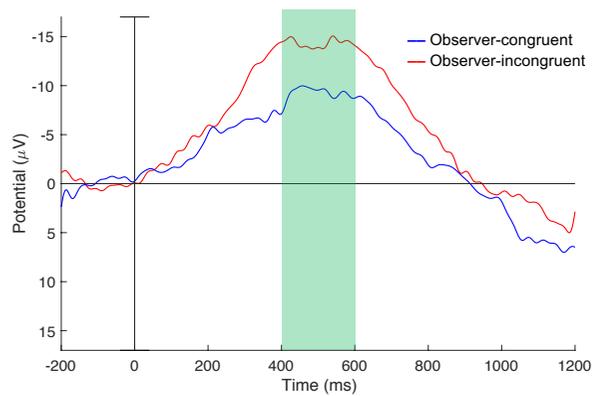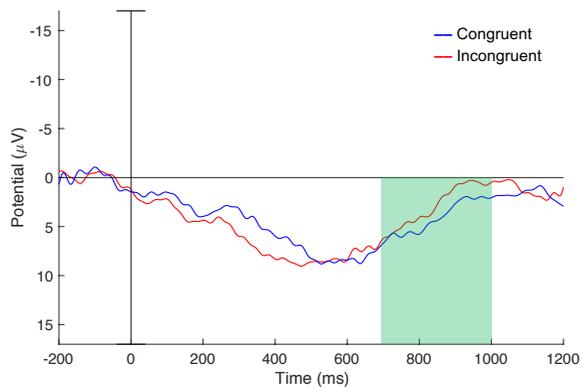
Experiment 2
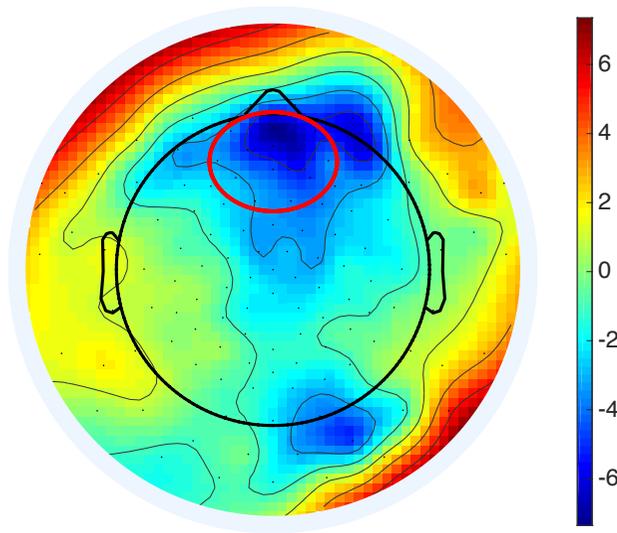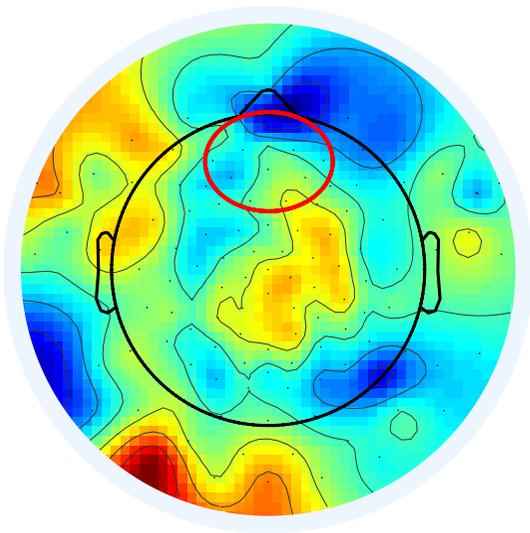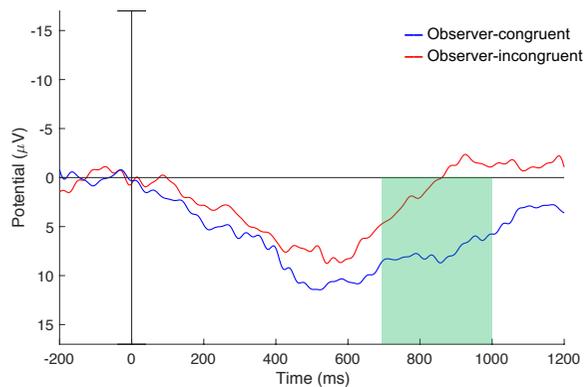
Fourteen-Month-Old Infants Track the Language Comprehension of Communicative Partners

Supplementary Material

SI Procedure

The full list of labels that were presented to the infants is show below. One item from the set of Parise and Csibra (2012), "teddy" ("nounours" in French) was replaced with "nursing bottle" ("biberon" in French), because "nounours" is an ambiguous word in French, referring not only to teddy bear, but also to toy animals in general.

French words	English translation
Balle	Ball
Banane	Banana
Biberon	Nursing bottle
Canard	Duck
Chaton	Cat
Chaussette	Socks
Chaussure	Shoe
Chien	Dog
Cuillère	Spoon
Lapin	Bunny
Livre	Book
Pomme	Apple
Tasse	Cup
Téléphone	Phone
Voiture	Car

SI EEG recordings

EEG Acquisition

High-density continuous EEG was recorded using Hydrocel Geodesic Sensor Nets (Electrical Geodesics Inc., Eugene, OR, USA), with 124 or 128 scalp locations referenced to the vertex (Cz). The common ground electrode was at the posterior location (between Cz and Pz),

standard for the net. The EEG signal was acquired using Electrical Geodesics Inc.'s Net Amps 400 amplifier, with a sampling rate of 500 Hz, using an anti-alias low-pass filter of 200 Hz.

EEG data processing

EEG data were analyzed using Net Station 4.5.6 (Electrical Geodesics Inc., Eugene, OR, USA). A band-pass filter between 0.3-30 Hz was applied to the raw EEG signal offline. The signal was segmented into to two categories (Congruent and Incongruent in Experiment 1, Observer-congruent and Observer-incongruent in Experiment 2), where each ERP epoch started 200 ms before and 1200 ms after the onset of the auditory stimulus. Next, the following automatic artifact rejection algorithms were applied to the data: A channel was automatically marked bad and rejected if the difference between the minimum and maximum amplitude value within the segment, in an 80 ms moving window, exceeded 200 μV at non-eye channels, 140 μV at vertical eye channels, indicating blinks, and 55 μV at horizontal eye channels, indicating horizontal eye-movements. The whole epoch was marked bad if more than 13 channels (10%) were marked bad. To ensure that infants paid attention at the time of playback, were engaged sufficiently to recognize the false belief of the Observer, and that the automatic algorithms picked up eye- and body-movements correctly, each trial was double checked using the video recording of the infants, and additional channels and segments were rejected manually if necessary. As a next step, bad channels were replaced using the interpolation algorithm of Net Station; then segments were averaged per condition; a baseline correction was applied using the 200 ms period preceding stimulus onset; and finally ERPs were re-referenced to the average reference.

Infants included in the final data analysis of Experiment 1 contributed, on average, 15.2 trials to the Congruent condition (SD = 5.6; range: 10-32), and 14.7 good trials to the Incongruent condition (SD = 4.31; range: 10-25). In Experiment 2, the average number of trials were 14.3 in the Observer-congruent condition (SD = 3.83; range: 10-22), and 14.2 in the Observer-incongruent condition (SD = 3.78; range: 10-21).

Based on previous findings (Friedrich & Friederici, 2005, 2008, 2010) and on visual inspection, the parietal region-of-interest (ROI) contained the following 13 electrodes: 62, 65, 66, 67, 70, 71, 72, 75, 76, 77, 83, 84, 90; while the frontal ROI the following 13 electrodes: 3, 4, 5, 9, 10, 11, 12, 15, 16, 18, 19, 22, 23.

SI Results

To test whether the frontal negativity is specific to Experiment 2, we run a 2x2 ANOVA in the 700-1000 ms time window over the frontal ROI with Condition (congruent, incongruent) as within-subject and Experiment (1 & 2) as between-subject factor. The main effect of Condition was significant, $F(1, 34) = 10.8$, $p = .002$, $\eta_p^2 = .24$, with the Incongruent condition being more negative than the Congruent condition. Importantly however, there was a trending (Bonferroni-corrected) Condition x Experiment interaction as well, $F(1, 34) = 5.18$, $p = .029$, $\eta_p^2 = .13$, which was explained by planned comparisons showing a significant effect only in Experiment 2, $t(17) = -3.42$, $p = .003$, Hedges's $g_{av} = .92$, but not in Experiment 1, $t(17) = -.87$, $p = .4$, Hedges's $g_{av} = .15$.

SI Figures
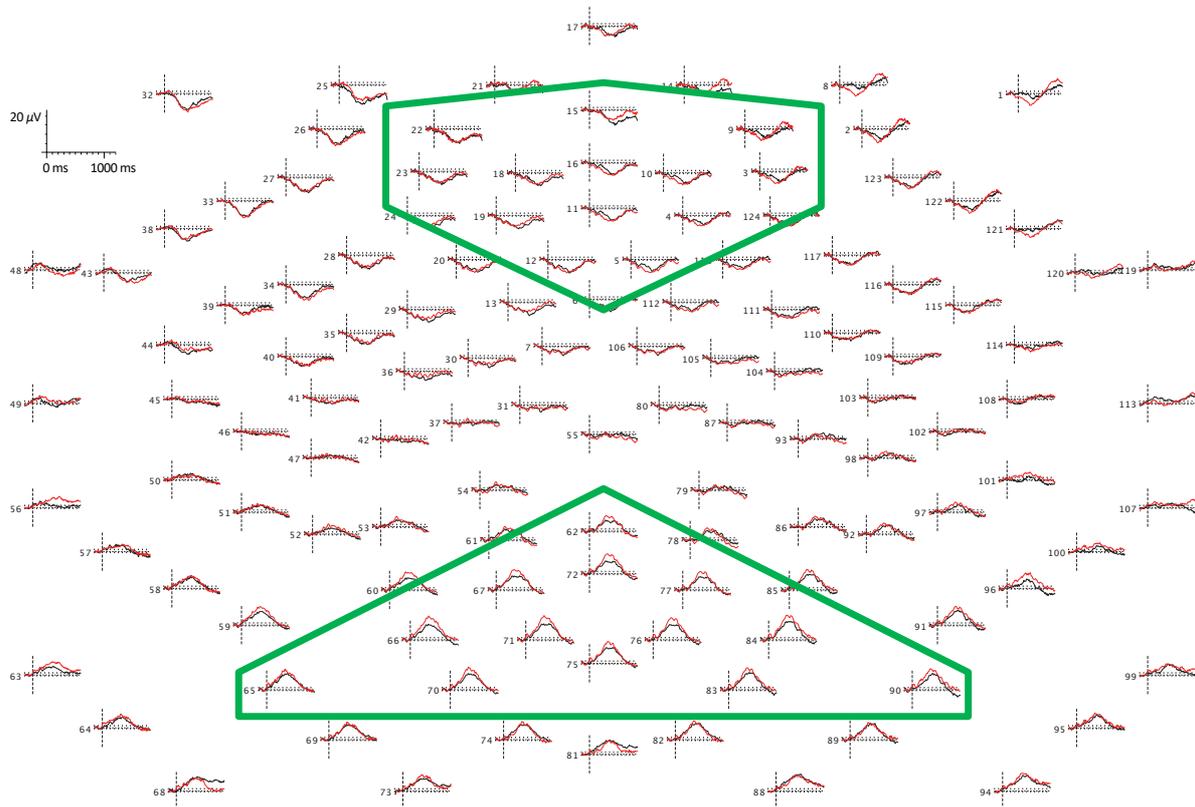
Figure S1. Grand-average ERP plots of Experiment 1 for all electrode sites. Negative is plotted upwards; frontal sites are above, parietal sites are below. Black line is the Congruent, red line is the Incongruent condition; green shapes indicate frontal and parietal ROIs. A typical infant N400 is apparent over posterior sites between 400-600 ms, but no frontal negativity can be observed.
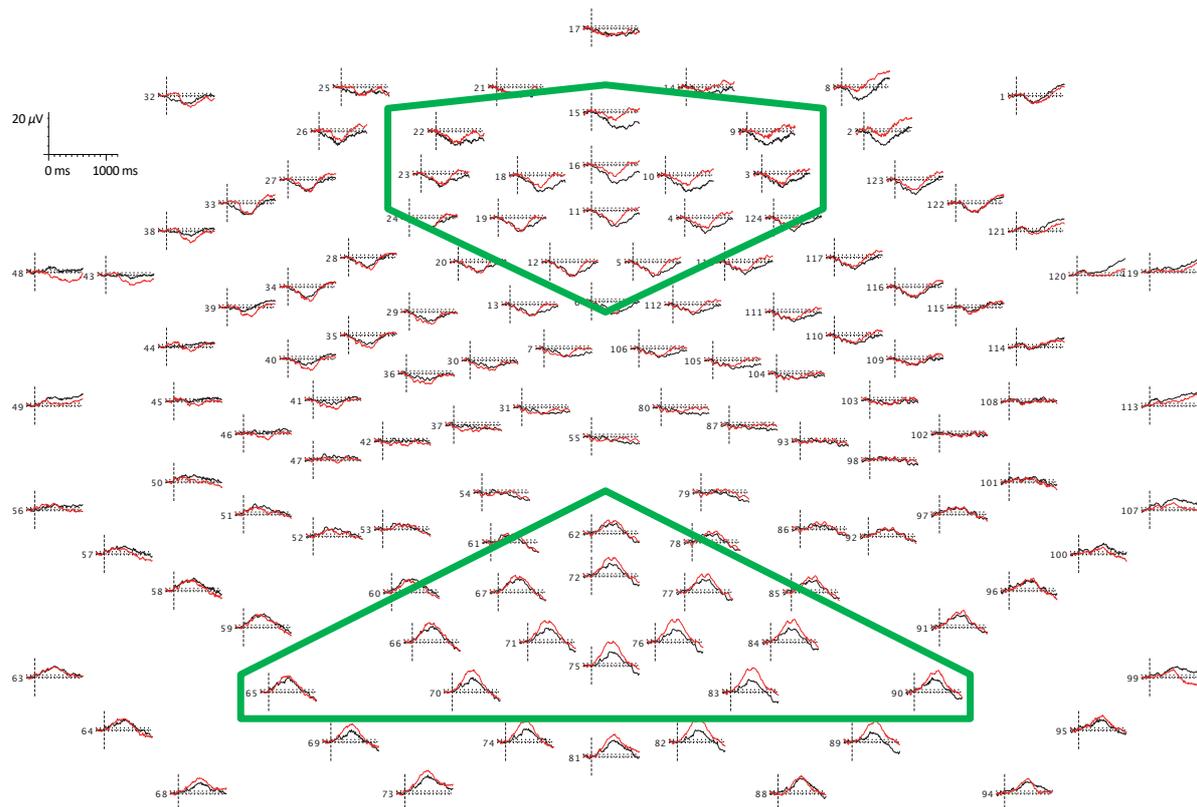
Figure S2. Grand-average ERP plots of Experiment 2 for all electrode sites. Negative is plotted upwards; frontal sites are above, parietal sites are below. Black line is the Observer-congruent, red line is the Observer-incongruent condition; green shapes indicate frontal and parietal ROIs. An N400-like effect can be observed over parietal electrode sites between 400-600 ms, and a late negativity is apparent over frontal sites between 700-1000 ms.