Electrophysiological evidence of early word learning

Caroline Junge [a,b,*], Anne Cutler [a,b,c], Peter Hagoort [a,b]

[a] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[b] Donders Centre for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
[c] MARCS Auditory Laboratories, University of Western Sydney, Australia

ABSTRACT

Around their first birthday infants begin to talk, yet they comprehend words long before. This study investigated the event-related potentials (ERP) responses of nine-month-olds on basic level picture–word pairings. After a familiarization phase of six picture–word pairings per semantic category, comprehension for novel exemplars was tested in a picture–word matching paradigm. ERPs time-locked to pictures elicited a modulation of the Negative Central (Nc) component, associated with visual attention and recognition. It was attenuated by category repetition as well as by the type–token ratio of picture context. ERPs time-locked to words in the training phase became more negative with repetition (N300–600), but there was no influence of picture type–token ratio, suggesting that infants have identified the concept of each picture before a word was presented. Results from the test phase provided clear support that infants integrated word meanings with (novel) picture context. Here, infants showed different ERP responses for words that did or did not align with the picture context: a phonological mismatch (N200) and a semantic mismatch (N400). Together, results were informative of visual categorization, word recognition and word-to-world-mappings, all three crucial processes for vocabulary construction.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The ability to learn names for things is an important milestone in language development. Around their first birthday infants start producing their first words. Early word learning, the first 50 productive words, is characterized as a slow and laborious process, with one or two newly produced words per week (e.g., Carey, 1978; Hollich, Hirsh-Pasek, & Golinkoff, 2000). Infants' first words typically refer to known exemplars or persons of high relevance to the individual child, such as 'mommy', 'hand' or 'dog' (Clark, 1993; Fenson et al., 1994). Around 18 months, infants' productive vocabulary dramatically increases, a phenomenon also known as the 'vocabulary spurt'. Note, however, that the vocabulary spurt is defined by changes in word production, not in word comprehension (Nazzi & Bertoncini, 2003). The average American 12-month-old might only produce six words, but can understand about 75 words (Bates, Dale, & Thal, 1995; Fenson et al., 1994; Golinkoff & Hirsh-Pasek, 2006). Thus vocabulary construction has started well before it is so evident in production. Vocabulary construction requires infants to show at least three skills: identifying a concept (categorization), identifying a word (recognizing words in a speech stream), and mapping words to objects or events (word-to-world mapping; Waxman & Lidz, 2002).

Although recent research has made good progress in studying infants' abilities to categorize speech or objects in isolation (cf. Waxman & Lidz, 2002), little is known about infants' ability to form their first word-to-world mappings, or about the interplay between the three vital vocabulary construction processes. There have been only a few experimental studies testing lexical-semantic knowledge of infants younger than 12 months, using the intermodal preferential-looking paradigm (IPL) (Bergelson & Swingley, 2012; Schafer, 2005; Tincoff & Jusczyk, 1999, 2000). In the IPL, infants typically see two objects, while the auditory label matches only one of the two. Their eye movements generally reveal a preference for (in the form of more looks to) the named versus the not-named referent. Using this paradigm, Tincoff and Jusczyk (1999, 2000) were the first to show that 6-month-olds already had some word–world associations. However, whether or not infants can recognize early words also depends on the experience they have had with them. Schafer (2005) compared infants' ability to map words to novel exemplars of common objects at 12 months when they had or had not received weekly training at home on a set of these objects from 9 months on. Only infants with this training experience looked longer at correctly named category exemplars, and only for the trained categories. With exemplars of un-trained categories, all infants performed at chance.

Thus training infants at 9 months boosts subsequent comprehension for the trained words at 12 months. But what happens in the infant brain during such learning? We here report an

* Correspondence to: Weesperplein 4, 1018 XA, Amsterdam, the Netherlands.
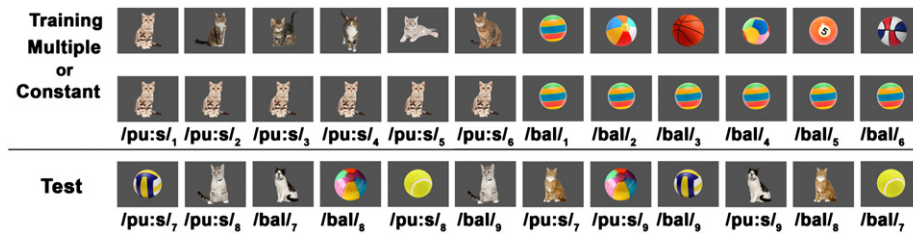E-mail address: Caroline.Junge@upcmail.nl (C. Junge).

Fig. 1. An example of the training and test phase for the block 'ball'–'cat'. Half of the infants saw this training phase with constant pairings (same picture six times); others saw this phase with multiple pairings: six different cats and six different balls. Each time a picture was presented, a novel token of a matching word was presented 1 s later, with the picture still on the screen. The test phase always consisted of three novel picture and word exemplars per category, each presented once in a congruous and once in a incongruous pairing.

exploration of neurophysiological changes over the course of typical early word-to-world pairings, using event-related brain potentials (ERPs). We tested 9 month-olds, the age at which infants in the Schafer (2005) study started their training. Infants were presented with ten training-test blocks, with each block contrasting two different real words that are typically acquired very early (e.g., 'ball'–'cat'). See Fig. 1 for an example of a training-test block for the example 'ball'–'cat'.

Each training phase showed six picture–word pairings of one semantic category (e.g. 'cat'), followed by six pairings of another category (e.g. 'ball'). We further manipulated the picture context of the training phase, specifically the type–token ratio of the presented categories. Behavioral research on visual categorization (e.g., Reznick & Kagan, 1983) has shown that infants familiarized with different exemplars show more preference for an unfamiliarized category at test (i.e., more habituation to the familiarized category) than infants presented with identical exemplars. In half of the training blocks in our study, pairings were constant (for each category, the same picture was presented six times), while the other blocks consisted of multiple pairings (six different pictures per category). After each training block we tested word comprehension of these two categories (e.g. 'ball' and 'cat'). Regardless of training context, the test phase always consisted of novel exemplars of trained categories, whose names were once congruous and once incongruous (from the contrasting category) to this novel exemplar. We measured ERPs time-locked to the pictures as well as to the words (which were presented one second after onset of picture).

Visual categorization researchers have reported a modulation of the Negative Central (Nc) component (Courchesne, Ganz, & Norcia, 1981; Nelson, 1994; cf. de Haan, 2007): a fronto-central negative deflection elicited for all types of visual stimuli, but generally largest for novel stimuli, peaking around 400–600 ms for 6-to-12-month-olds. Its amplitude is considered to be an index of attention, or of recognition memory. In object categorization studies, the Nc is associated with preference for one (novel) category over another (familiarized category) (Grossmann, Gliga, Johnson, & Mareschal, 2009; Jeschonek, Marinovic, Hoehl, Elsner, & Pauen, 2010; Quinn, Doran, Reiss, & Hoffmann, 2010; Quinn, Westerlund, & Nelson, 2006). Hence, we hypothesized that for ERPs time-locked to pictures in the training phase, the Nc is reduced for the second half of the training phase (repetition effect). The type of pairings should also influence the Nc: We expected the amplitude to be larger for different exemplars than for repetitions of the same exemplars.

Studies on unimodal auditory word processing in infants, on the other hand, show that ERPs are more negative for familiar words than for unfamiliar words (N200–500, often largest on frontal electrodes). This is the case for known versus unknown single words (e.g., Thierry, Vihman, & Roberts, 2003) as well as for familiarized versus unfamiliarized low-frequency word forms in continuous speech (e.g., Kooijman, Hagoort, & Cutler, 2005; Goyet, de Schonen, & Nazzi, 2010). The N200–500 has also been reported in studies of novel word-to-world mappings in 6- and 14-month-old infants: ERPs time-locked to the onset of words revealed that the more often words were consistently presented with the same novel objects, the larger the N200–500 (Friedrich & Friederici, 2008, 2011; Torkildsen et al., 2009; note however that for 6-month-olds in the 2011 study this effect appeared not only when words were consistently paired with novel objects, but also when novel words were randomly paired with objects, so that it could have been evoked by auditory word repetition rather than by word–object integration). Besides the N200–500, Friedrich and Friederici (2008, 2011) also observed for parietal electrodes a decrease in negativity for the more familiarized words, which they interpreted as semantic priming of the picture context. Extrapolating from these findings to the present study, we hypothesized that ERPs time-locked to words in the familiarization phase would elicit an increased N200–500 for the second half compared to the first half of a training phase (repetition effect). This word familiarity effect could be influenced by our picture context manipulation: if identifying a category is more difficult given different training pictures than given the same picture repeatedly presented, the N200–500 could be decreased or delayed.

In the test phase of our study, we used the picture–word paradigm to test word-to-world mappings for exemplars infants had neither seen nor heard before. In adults, real words matching a picture elicit a smaller N400 than non-matching words (e.g., Desroches, Newman, & Joanisse, 2009; Friedrich & Friederici, 2004). The N400, a negative component peaking around 400 ms after stimulus onset at posterior electrodes, is observed in children as well as in adults and provides a reliable index of lexical-semantic processing, specifically the difficulty of accessing and integrating a word into its current context (Kutas & Hillyard, 1980; cf. Kutas & Federmeier, 2011). The youngest age for which the N400 has been reported is 6 months (Friedrich & Friederici, 2011).

Results from studies using picture–word matching (without prior familiarization) to test lexical-semantic processing in infants have however been inconsistent with regard to the N400. Mills and colleagues (e.g., Mills, Conboy, & Paton, 2005) found a comparable N400 latency across development, from before the 'vocabulary spurt' (at 13 months), through 20 months, to 3 years, and in adulthood. Friedrich and Friederici (2004, 2005a, 2005b), on the other hand, found no N400 for 12-month-olds, but an N400 in 14- and 19-month-olds that was delayed compared to the adult response. Since, regardless of whether the N400 was present, infants at all ages showed the N200–500 described above, Friedrich and Friederici interpreted the N200–500 as a "phonological-lexical priming effect", and interpreted absence of an N400 as evidence that words were not yet processed to a level which would support semantic priming. They also observed that the N400 was linked to word learning: 12-month-olds who produced

more than four words showed an N400, but 12-month-olds with smaller vocabularies did not (even for words that parents had rated as understood by their infant; Friedrich & Friederici, 2010). Parise and Csibra (2012) showed that even in 9-month-olds, incongruous word–object pairings can elicit an N400, but in this case the effect only appeared when the critical words were uttered by a familiar voice (the mother's), not by the experimenter. The presence of an N400 in infancy may thus depend on multiple factors, including task characteristics as well as brain maturation. Though the conditions for appearance of an infant N400 may be not yet fully understood, the combined results do suggest that when it appears, it can be taken as a neurophysiological marker of successful semantic processing of real words.

However, there is also clear indication that increasing infants' familiarity with the experimentally tested words, either by training on certain word–object pairings (Schafer, 2005) or by using a familiar speaker's voice (Parise & Csibra, 2012), will raise the likelihood of N400 appearance. The training phase in our study should provide the requisite conditions for this. The training we used was modeled on that used in novel-word learning studies. Test phases of such studies have shown a larger N400 for incongruous words, both immediately after familiarization (Torkildsen et al., 2008) or a day later (only for 14-month-olds, not for 6-month-olds; Friedrich & Friederici, 2008, 2011). Note that novel word learning, in which infants learn labels for objects with which they have no prior experience, differs from early word learning in real life in that in the latter case, infants have seen objects multiple times before they learn their labels. In our test phase infants always saw novel exemplars of trained (and known) categories, and heard novel word tokens, either congruous or incongruous. We predicted an N400 for incongruous pairings, indicating that infants can attach labels to novel exemplars. We further predicted that infants would find it easier to recognize novel tokens as belonging to a certain category when they had been familiarized with multiple tokens in the training phase. This would then be reflected in a larger N400 for incongruous words for the multiple pairing condition than for the constant pairing condition. Together, the training and test phases offer an window on the neural mechanisms of early word learning.

2. Methods

2.1. Participants

Twenty nine-month-old infants (nine female) participated. Their mean age was 282 days (SD=6.1 days). An additional 11 infants were tested, but excluded due to inattentiveness (n=2); refusal to wear the cap (n=1); computer problems (n=1); or retaining too few artifact-free trials (n=7). All subjects were healthy, full-term infants from monolingual Dutch families with no history of neurological or language impairments. The majority had college-educated parents. Infants were recruited from the Nijmegen Baby Research Center Database. Parents signed informed consent forms, and received 20 euro and a photograph of their child taken after the experiment in appreciation of their participation.

2.2. Materials

Twenty easily depicted basic-level nouns were selected from the Dutch version of the MacArthur–Bates Communicative Development Inventory (CDI: Fenson et al., 1993; N-CDI: Zink & Lejaegere, 2002), based on likelihood of understanding the word at 12 months and of being visually familiar with its referent at 9 months. We selected words based on 41 N-CDIs for 12-month-olds (21 boys; mean age 366 days, SD=10.1 days) collected previously in the Nijmegen Baby Research Center. The mean percentage of 12-month-olds understanding the 20 selected words was 47.6% (SD=11.7%). The 20 words came from the following six semantic domains: animate (4); clothing (3); body parts (3); furniture (3); food (3); toys (2) and vehicles (2). See Table 1 for an overview of the words.

For each word, ten auditory tokens were recorded by a female native speaker of Dutch in a sound-proof booth in a lively child-directed manner, and digitized at a sampling rate of 44.1 kHz. The mean duration of words was 621 ms (SD=103 ms; mean SD per word=45 ms).

Per category, there were 10 different color stock photographs (modified with Adobe Photoshop). The photographs were roughly of the same size and appeared in isolation on a dark-grey background.

2.3. Procedure

The experiment comprised 240 picture–word trials. Pictures stayed on the screen for 2200 ms, with the word presented 1000 ms from picture onset. The inter-trial interval was 1000 ms. There were ten training-and-test blocks, with each block contrasting two semantic categories. These two categories always came from different semantic domains, e.g. 'cat' (animal) versus 'ball' (toys). Moreover, the labels of the two contrasted categories did not share any overlap in onset or vowels (see Table 1). A block consisted of 12 trials in the training phase, followed by 12 trials in the test phase. Each training phase started with six picture–word trials of one category, followed by six picture–word trials of the second category. The type of training (constant versus multiple pairings) varied as a within-subjects variable but was kept constant per block. Half of the blocks consisted of constant pairings; that is, the same picture was presented six times, each time with a different token of the congruent word. The other five blocks consisted of multiple pairings: six different exemplars of a category, each paired with a different auditory token of the correct verbal label. There were no more than two blocks with the same type of training in a row.

The test phase was alike for each training type: There were three novel pictures and words per category that had not been presented in the training phase. Each picture and each word were presented twice: once in the congruent condition and once in the incongruent condition. In this way congruency effects could not be due to physical differences between pictures or between words. Presentations of the six congruous and six incongruous picture–word pairs were quasi-randomized so that the same token never appeared consecutively, and with no more than two congruous or incongruous trials in a row.

To avoid item-specific and order effects, we constructed four presentation lists, counter-balancing type of pairing per block, order of blocks, and order of categories within a block. Five infants heard each list. Each list was randomized, so that for some infants certain items were presented in the training phase that were part of the test phase for others with the same list.

Table 1
The 20 nouns used in the experiment, split by block. Per block, two words were contrasted; for instance, 'cat' and 'ball' for the first block. For each word, its phonetic transcription in Dutch, category membership (domain) and the mean average ratings of visual familiarity and likelihood word comprehension are reported. Parents rated visual familiarity as how often their child would see each semantic category in real life or in books on a scale from 0 (never)–5 (every day), with 3 as 'once a week'. They rated word comprehension on a scale from 1 (not)–5 (well), with 3 as 'maybe'.

Block	Semantic category A	Phonetic transcription	Domain	Visual rating	Word rating	Semantic category B	Phonetic transcription	Domain	Visual rating	Word rating
1	**Cat**	[pʊːs]	Animate	2.3	2.6	**Ball**	[bɑl]	Toy	4.5	3.7
2	**Mouth**	[mɔnt]	Body part	4.6	3.1	**Chair**	[stʊːl]	Furniture	4.6	3.1
3	**Car**	[ɔuto:]	Vehicle	4.6	3.4	**Banana**	[bɑnaːn]	Food	3.2	2.9
4	**Foot**	[vʊːt]	Body part	4.6	3.4	**Bed**	[bɛt]	Furniture	5.0	3.9
5	**Dog**	[hɔnt]	Animate	2.2	3.2	**Book**	[bʊːk]	Toy	4.5	3.7
6	**Sock**	[sɔk]	Clothing	4.3	3.0	**Bottle**	[flɛs]/[flɛsje]	Food	4.9	4.3
7	**Baby**	[beːbiː]	Animate	3.6	3.2	**Cookie**	[kʊːkje]	Food	2.8	2.8
8	**Coat**	[jɑs]	Clothing	4.4	2.9	**Bicycle**	[fiːts]	Vehicle	3.3	2.7
9	**Bath**	[bɑt]	Furniture	3.8	3.6	**Cow**	[kʊː]	Animate	2.0	2.0
10	**Shoe**	[sxʊːn]	Clothing	4.0	2.8	**Hand**	[hɑnt]	Body part	5.0	2.6

During the experiment the infant sat in a child seat in a sound-proof booth. Visual stimuli were presented on a 19 in. computer screen with a 60 Hz refresh rate and a 1024 × 768 pixel resolution, situated 1 m in front of the child. Words were presented at an intensity of 65 dB through two loudspeakers placed 1.5 m in front of the child. Infants were video-monitored to control whether they processed the visual information. Attention-grabbers (short video clips of cartoons, e.g. a moving merry-go-round or a duck going to bed) were played after every two blocks and whenever infants were losing interest. A parent sat next to the child, listening to masking music through closed-ear headphones. Breaks were taken when necessary. The experiment lasted 14 min, and a whole session about an hour.

After the experiment, parents filled in the N-CDI for infants. Infants produced between zero and two words (mean .15, SD .49) from the N-CDI, and understood on average 53 words and utterances (SD 41.7). Parents also filled in two questionnaires designed for this experiment. These measured visual familiarity and word comprehension for each of the 20 categories used in the experiment, with a higher rating indicative of a better understanding of a category. Average visual familiarity, rated on a scale from 0 to 5, was 3.9 (SD .48), corresponding to seeing items at least once a week. Average word comprehension, on a scale from 1 to 5, was 3.1 (SD .56), corresponding to 'maybe'. There were no correlations between subject mean ratings and vocabulary scores (Pearson's $r = +.02$ to $+.37$; $p > .11$). When mean ratings per word were calculated (averaged over infants), we observed a significant positive correlation between visual and word familiarity ratings, indicating that the more often objects were seen, the better their labels were understood ($R = +.61$, $p = .005$).

2.4. EEG recordings and pre-processing

EEG was recorded with a sampling rate of 500 Hz, using an infant-size BrainCap with 24 inserted Ag/AgCl sintered ring electrodes, placed according to the extended 10–20 system (F7, F3,Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, Oz). Vertical eye movements and blinks were monitored via a supra- to sub-orbital bipolar montage and horizontal eye movements via a right-to-left canthal bipolar montage. Electrodes were referenced online to the left mastoid and re-referenced to linked mastoids offline. Impedances were kept below 5 kΩ for the ground and reference electrodes, and below 20 kΩ for the remaining electrodes. The signal was filtered with an on-line filter of .01–200 Hz and an off-line filter of .1–30 Hz. Trials were time-locked to the onset of pictures as well as to the onset of words. Based on video recordings we rejected trials (for both words and pictures) when infants were not looking at the screen.[1] Individual trials with a baseline of 200 ms were furthermore screened for artifact from 200 ms before to 1000 ms after target onset. They were automatically rejected when amplitudes exceeded ± 200 μV, and manually rejected when we detected clear correlations with the eye channels or with activity in the right mastoid during recording. The electrode Oz was excluded from analysis due to excessive artifact. The persons coding the video-recordings and performing the manual artifact rejection of the remaining trials were blind to the conditions of the experiment. For each infant, we calculated average waveforms per condition, with a minimum of ten artifact-free trials per condition. Infants attributed on average 15.7 artifact-free trials per condition (SD 3.0).

2.5. Statistical analyses

We report both behavioral and electrophysiological results. For behavioral results we analyzed the number of trials when infants looked away. For electrophysiological results, we calculated mean amplitudes for selected time windows per condition for each of the 20 lateral electrodes. Time windows were selected based on visual inspection of the waveforms. For all ANOVA tests, we used the Huynh–Feldt epsilon correction and we report here original degrees of freedom, adjusted p-values, and adjusted effect sizes (partial eta-squared: η^2). To test topographic distribution of the relevant effects, we added the factors anterior/ posterior (2), hemisphere (2: left, right), and electrode (5) to the ANOVAs. This created four quadrants of the brain: left frontal (F7, F3, FC1, FC3, C3); right frontal (F8, F4, FC2, FC6, C4); left posterior (T7, CP5, CP1, P7, P3), and right posterior (T8, CP6, CP2, P8, P4).

To assess the effect of repetition in the training phase, we compared results for the first (first three trials) versus the second part of the training phase (second three trials), with type of pairing (constant versus multiple pairing) as a within-subjects factor. For the training phase, we analyzed ERP repetition (2) and type of pairing (2) effects separately for visual processing (time-locked to the onset of picture), and for word processing (time-locked to the onset of the word).

For the test phase, we compared results for congruous versus incongruous words, with again type of pairing (in training phase) as a within-subjects factor. We only analyzed selected time windows time-locked to the word. As all words

and pictures in the test phase were presented once in the congruous and once in the incongruous condition, a congruity effect could therefore only be due to the pairing between pictures and words.

3. Results

3.1. Behavioral results

First we calculated the number of trials that infants were not looking at the screen from the first and second part of the training phase, separately for the two types of pairings (i.e., constant and multiple pairings). The maximum number of picture–word trials that infants could observe per condition was 30. For the first versus second part of the training phase, infants did not look at the screen for 2.35 (SD 2.98) and 3.35 (SD 3.07) trials for blocks with constant pairings, and 2.10 (SD 2.85) and 2.70 (SD 3.08) for blocks with multiple pairings. There is a marginal effect of repetition ($F_{1,19} = 3.67$, $p = .070$, $\eta^2 = .16$), indicating that infants were more likely to lose interest in the second block of training. There is no main effect or interaction with type of pairing ($F_{1,19} < 1.64$, $p > .20$).

For congruous versus incongruous trials in the test phase, infants did not look at the screen for 4.10 (SD 4.39) and 3.95 (SD 4.14) trials for the constant pairings, and 3.95 trials (SD 3.89) and 3.70 (SD 4.09) for the multiple pairings. There are no significant effects for the test phase ($F_{1,19} < 1$). Infants were, however, significantly more likely to divert their gaze in the test phase than in the training phase ($F_{1,19} = 6.15$, $p = .023$, $\eta^2 = .24$). The type of pairings they received in the training phase did not influence this difference ($F_{1,19} < 1$).

3.2. Electrophysiological results

3.2.1. Training phase, time-locked to picture onset

Fig. 2 shows the grand average waveforms time-locked to the onset of pictures, with a 200 ms baseline, up to the onset of the word at 1000 ms, for the four conditions in the training phase (first versus second half of the training phase, for constant and multiple pairings, respectively). All conditions elicited a large broadly-distributed negative wave from 300 ms onwards, which is typical for visual processing (de Haan, 2007). Only on lateral parietal electrodes (P7, P8) do we observe an opposite polarity. Based on visual inspection, we chose the 300–750 ms interval for analysis, which is a standard interval for the Nc-component (de Haan, 2007). Statistical analyses show that there are main effects of repetition ($F_{1,19} = 12.2$, $p = .002$, $\eta^2 = .40$) and of pairing type ($F_{1,19} = 8.21$, $p = .010$, $\eta^2 = .30$), but no interaction between the two ($F_{1,19} = 1.76$, $p = .20$, $\eta^2 = .09$). First, the Nc is reduced for the second part compared to the first part of the training phase, regardless of type of pairing. Second, the Nc is more reduced for constant than for multiple pairings in each of the parts of the training phase. In addition, both repetition and type of pairing have an interaction with posterior/anterior distribution ($F_{1,19} = 12.0$, $p = .003$, $\eta^2 = .39$; $F_{1,19} = 5.08$, $p = .036$, $\eta^2 = .21$, respectively), indicating that both the repetition and the type of pairing effects are largest over anterior electrodes.

3.2.2. Training phase, time-locked to word onset

We further examined ERPs from word onset. Fig. 3 shows the grand average waveforms time-locked to the onset of the word, again for the four conditions in the training phase (first versus second half of the training phase, for constant and multiple pairings, respectively). Whereas we observed a large negative wave peaking around 600 ms for picture processing, we see a large positive wave peaking around 400 ms on anterior electrodes

---

[1] For four infants we did not have video recordings. Here, we used information from additional attention-grabbers as a measure of inattentiveness, and rejected the two trials prior to the onset of these attention-grabbers.
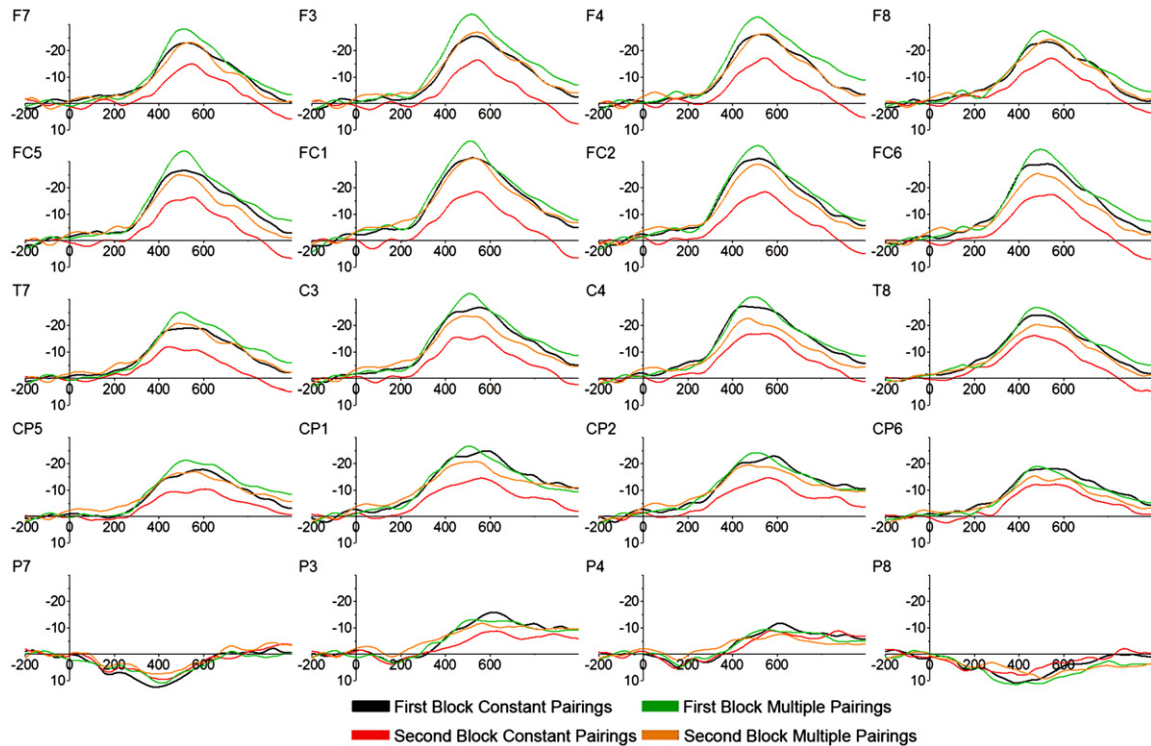
Fig. 2. Grand average waveforms, time-locked to picture onset, for the first and second block of the training phase, for the constant and multiple pairings. (In this and all following ERP figures, electrodes are arrayed for most anterior (top) to most posterior (bottom), and from left to right as they were positioned on the scalp; negativity is plotted upwards; an 8 Hz. low-pass filter has been applied for illustrative purposes only).
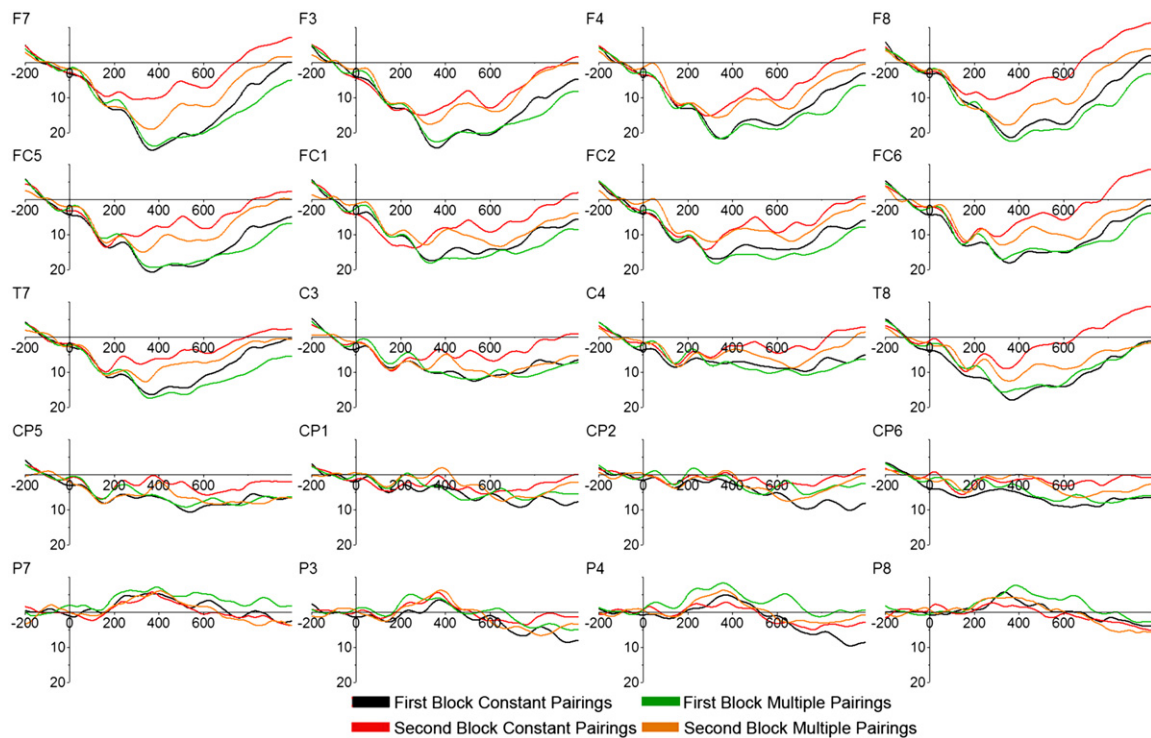

Fig. 3. Grand average waveforms, time-locked to word onset, for the first and second block of the training phase, for the constant and multiple pairings.

for auditory word processing. Other studies of auditory single word processing have also reported a large positive wave (e.g., Kooijman et al., 2005; Sheehan & Mills, 2008). This makes it more likely that this component is reflective of word processing rather than being a late component of picture processing. As predicted, this positive wave is reduced by repetition.

Visual inspection shows that the point of time where the ERPs for the first and second block diverge is around 250–300 ms, and
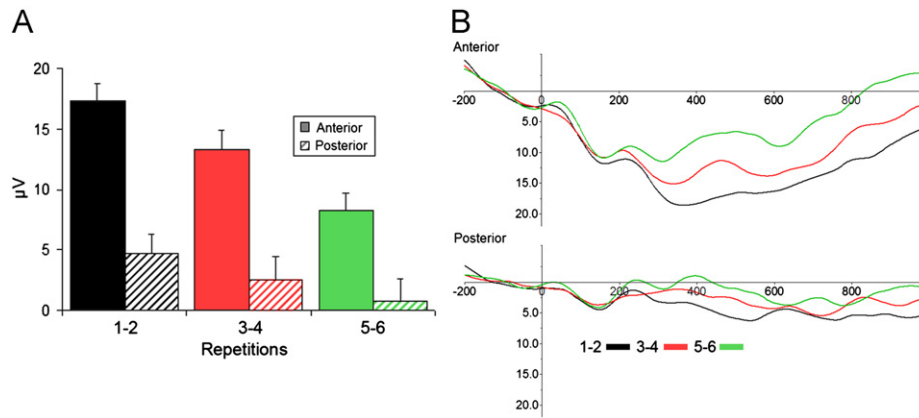
Fig. 4. The ERP for word repetition is graded and becomes less positive the more often a word is presented. ERPs are here time-locked to the onset of the word, averaged separately for anterior (average of F7, F3, F4, F8, FC5, FC3, FC2, FC6, C3 and C4) and for posterior (average of T7, T8, CP1, CP5, CP2, CP6, P3, P7, P4, P8) electrodes: (A) for the time window 300–600 ms, with error bars reflecting one standard error of the mean. (B) The time course of the word recognition effect, on pooled anterior and pooled posterior electrodes, respectively.

that this is the same for both constant and multiple pairings. Indeed, statistical analyses for the 300–600 ms interval show a main effect of repetition ($F_{1,19}=36.4$, $p < .001$, $\eta^2=.66$), but no main effect of or interaction with type of pairing ($F_{1,19}=.23$, $p=.64$, $\eta^2=.01$; $F_{1,19}=1.48$, $p=.24$, $\eta^2=.07$, respectively), or any other interactions involving distribution and type of pairing. The only significant interaction is between repetition and anterior/posterior ($F_{1,19}=42.1$, $p < .001$, $\eta^2=.69$): The effect of repetition is largest over anterior electrodes.

Since there is no influence of type of pairing on word repetition, we collapsed over trials with multiple and constant pairings. To investigate the effect of word recognition in a graded manner, we then compared ERPs for word onset in the training phase for picture–word combinations 1–2 and 3–4 and 5–6. There is again a main effect of word repetition ($F_{2,38}=10.9$, $p < .001$, $\eta^2=.37$), and an interaction of word repetition by anterior/posterior ($F_{2,38}=14.4$, $p < .001$, $\eta^2=.43$). Fig. 4 demonstrates this: The more often a word is presented, the more reduced the ERP becomes. The graded familiarity effect is more pronounced over anterior electrodes (Fig. 4A), and starts around 300 ms (see Fig. 4B for the time course of word recognition).

3.2.3. Test phase, time-locked to word onset

Fig. 5 shows the grand average waveforms time-locked to the onset of the word, for the congruous and incongruous conditions, split by type of pairing. As in the training phase, we observe a large positive wave which we associate with auditory word processing. There are two time windows where the ERPs for congruous words diverge from those for incongruous words. From 200 to 300 ms (N2 window), incongruous words elicited a larger negative going deflection than congruous words, which is more pronounced over anterior electrodes. There is a main effect of congruity ($F_{1,19}=5.64$, $p=.028$, $\eta^2=.23$), but no main effect of or interaction with type of pairing ($F_{1,19} < 1$; $p > .5$), or any interactions with quadrants.

From 300 to 400 ms the waveforms converge. The effect of congruity is then no longer significant ($F_{1,19}=1.63$, $p=.22$, $\eta^2=.18$). The ERPs for congruous versus incongruous words diverge again from 400 to 600 ms, for constant as well as for multiple pairings. For this last time window the ERPs for incongruous words again show a more negative-going deflection than the ERPs for congruous words ($F_{1,19}=7.52$, $p=.013$, $\eta^2=.28$), but there are no interactions or main effects with type of pairing or with quadrant factors ($F_{1,19} < 2.45$, $p > .13$). Because there is no

interaction with pairing type, we could once more collapse over type of pairings. Here, the average N400 effect over posterior electrodes is significantly related to vocabulary size at 9 months ($R = -.48$, $p=.034$). Fig. 6 illustrates this: The larger the negativity over parietal electrodes, the more words the infant is reported to comprehend. To examine the interplay between previous item-specific knowledge and semantic processing, we also calculated whether there was a relationship between the size of the N400 and how well infants comprehended the specific words during test, on a scale from 1 to 5, as judged by their parents. Mean parental ratings on words used in experiment however were neither related to the size of posterior N400 effect ($R=+.08$, $p=.75$) nor to N-CDI scores on overall vocabulary ($R=+.37$, $p=.11$).

4. Discussion

The primary goal of this study was to obtain electrophysiological evidence relevant to early word learning in 9-month-old children. After a familiarization phase of six picture–word pairings per category, comprehension for novel exemplars was measured in a picture–word matching paradigm. Results gave evidence of visual categorization, word recognition and word-to-world-mappings, all three crucial processes for vocabulary construction. During the training phase, infants displayed a suppressed Nc effect for picture repetitions, which was also modulated by the type–token ratio of picture context. ERPs time-locked to words also showed a Word Familiarity effect, which was not affected by type of pairings. Results from the test phase provide clear support that infants integrated word meanings with (novel) picture context. Here, we observed an N200 and an N400 effect for words that are incongruous with the pictures. In the following sections we discuss these findings separately.

4.1. The Nc effect for picture processing

The Nc in this study was attenuated by repetition as well as by the type–token ratio of pictures. Traditionally, the amplitude of the Nc is considered to be an index of attention allocation (Nelson, 1994): Its amplitude can be influenced by novelty or by repetition priming (e.g., Quinn et al., 2006, 2010; Wiebe et al., 2006), as well as by saliency (i.e., larger for mother's face versus stranger's face; de Haan & Nelson, 1997). Behavioral measures in this study also revealed that infants progressively look less at
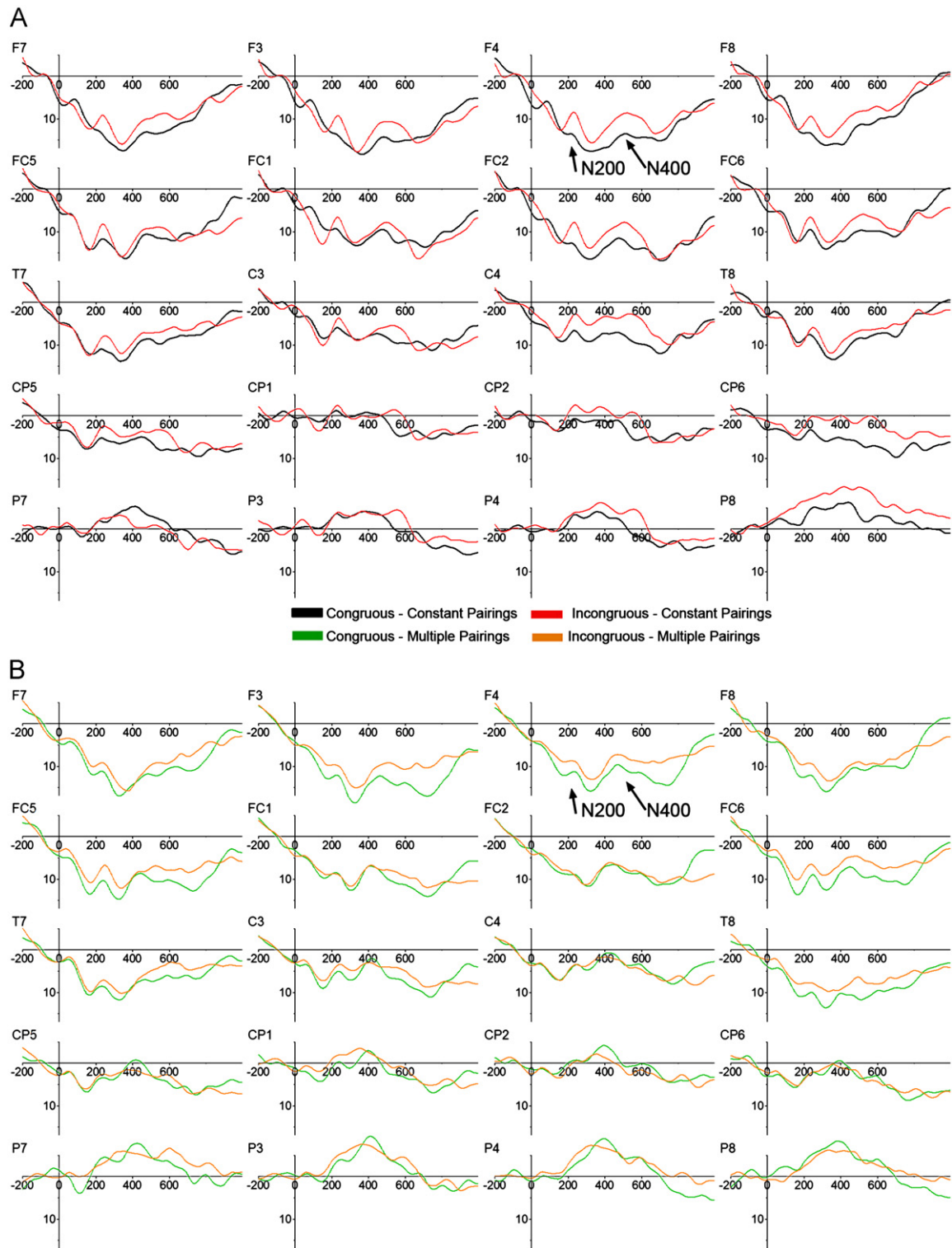
Fig. 5. Grand average waveforms, time-locked to word onset, for the congruous and incongruous pairings, for the constant (A) and multiple pairings (B), respectively.

exemplars of one category over time, which indicates that they habituated to the trained category. Hence, the behavioral results further provide evidence that infants were sensitive to repetition, suggesting that it is attention that modulated the Nc here. On the other hand, there was no behavioral evidence that infants were sensitive to the type–token ratio of pictures, although the Nc was also influenced by the type of pairings received in the training phase. Whether it was attention, recognition, or an interplay

between the two that drove the modulation of the Nc, its presence implies that the infants have encoded the pictures.

Other visual categorization studies with a familiarization phase of just one category, however, reported a modulation of the Nc only for the test phase, when it was increased for a novel category compared to the familiarized category (Grossmann et al., 2009; Quinn et al., 2006, 2010). Repetition was in that case studied by contrasting an ERP average of the first half (18–20
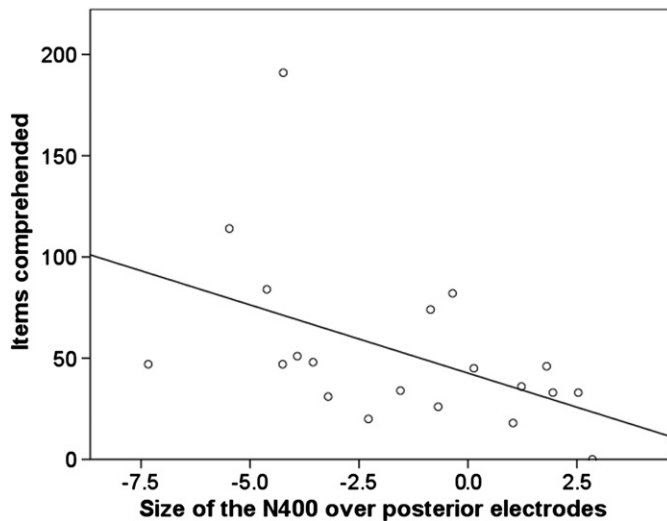
Fig. 6. A significant correlation between the size of the Congruency effect (incongruous–congruous) over posterior electrodes at the *X*-axis, and number of items (words and typical utterances) understood at nine months, at the *Y*-axis. This is still significant when excluding the outlier at $(-4,193; 187)$: $r = -.51$, $p = 028$.

tokens) versus the second half of the familiarization phase, whereas we compared the first three versus the last three tokens of a category. Consequently, the lack of a differentiation of the Nc for the familiarization phase in these studies could have been masked by averaging over too many successive repetitions of the same category. This is in line with recent research showing that the Nc is sensitive to saliency or novelty, which can change over the course of an experiment (Stets & Reid, 2010).

In sum, we have shown that when infants are familiarized with only six tokens of a category, both repetition and type–token ratio of exemplars influence a mid-latency component associated with attention and recognition. The timing of this effect suggests that infants have identified the picture before the presentation of the word.

4.2. The word familiarity effect

In a pictorial context, ERPs for words in the second half of the training phase became more negative than in the first half (N300–600). A similar negative middle-latency familiarity effect (i.e., N200–500) has been observed in two types of auditory word processing studies in infants: both for words rated by parents as known versus unknown to their child (e.g., Mills, Coffey-Corina, & Neville, 1993, 1997; Thierry et al., 2003), as well as for unknown but familiarized words versus unfamiliarized words (e.g., Kooijman et al., 2005). Although Mills and colleagues showed that the N200–500 for 20-month-olds is related to word meaning and not to word familiarity (Mills, Plunkett, Prat, & Schafer, 2005), Junge and colleagues hypothesized that for younger infants the same recognition mechanism is sensitive to word form repetition, so that meanings of words can be learned (Junge, Hagoort, Kooijman, & Cutler, 2010; Junge, Kooijman, Hagoort, & Cutler, 2012). Results from the present study confirm the suggestion that the same mechanism is involved in recognizing word familiarity as well as word meaning. The word familiarity effect in the current study was measured by repetitions for early typical words that the majority of 12-month-olds would understand. Infants in this study comprehended the words to some extent (i.e. average parental rating was 'possibly understood'), yet it was repetition of these items that elicited the word familiarity effect. Moreover, this effect was graded: The more often the item was repeated, the

more negative the corresponding ERP became. It had a similar polarity and distribution as observed in auditory familiarization studies, although its latency was delayed by 100 ms (i.e. N300–600 observed in a cross-modal context versus N200–500 in an auditory context). This delay could be the consequence of having words presented in the context of a picture.

The finding that the type–token ratio of pictures did not influence the word familiarity effect suggests that infants have encoded the picture before word onset. The timing of the visual repetition effects (300–750 ms from picture onset) also indicates that infants have encoded the picture before the word was presented (at 1000 ms after picture onset). However, an alternative hypothesis is that this word familiarity effect reflects purely auditory word repetition, regardless of the picture presented. Consider that 6-month-olds showed a word familiarity effect (N200–500 for repetitions) in a novel word learning situation where no mapping was possible: Novel words were randomly paired with novel objects (Friedrich & Friederici, 2011). Results from the test phase from that study showed, however, that the same infants continued to show the N200–500 response to words presented with the congruent novel rather than with an incongruent novel object, suggesting that N200–500 for infants this young is not only sensitive to word repetition, but also to the integration of word processing into a pictorial context. The increased N200–500 observed in our study could thus reflect a facilitation of word processing resulting from auditory word repetition, or from repeated pairing of a word with the same visual category, or from a combination of these two possibilities. The similarities between the word Familiarity effect in this and other auditory studies further suggest that infants recognize early typical words in a similar way with or without supporting context.

4.3. Semantic congruity effects

In the present study infants always saw three exemplars of each visual category that were not presented in the training phase. We found two semantic congruity effects, i.e., differences between trials when spoken words did versus did not fit the picture context. This indicates that these 9-month-olds were not only able to recognize novel visual tokens as belonging to the same types, but were also able to attach the correct labels to them. In other words, these infants show signs of lexical-semantic processing skill. This was the case when they were familiarized prior to test either with one or with six different visual tokens of a category. Furthermore, these congruity effects also suggest that the Word Familiarity effect in the training phase was not due to auditory repetition alone, but reflects word recognition in the context of visual information.

The first semantic congruity effect was observed in the time window 200–300 ms after word onset. A semantic congruity effect starting at 200 ms has also been observed in other infant studies using the match–mismatch paradigm; however, it is often considered to be part of the N400, i.e. N200–600 (Mills et al., 2005; Torkildsen et al., 2008). Indeed, in adult literature, the onset of the N400 is frequently observed from 200 to 600 ms, during which ERPs for match or mismatch conditions never converge. In our study, however, the two effects seem to occur in two separate latency windows, which suggests that the two effects might be reflecting different processes. An early congruency effect, separate from the N400, has also been reported for both 18-month-olds (Sheehan, Namy, & Mills, 2007) and 14-month-olds (Duta, Styles, & Plunkett, 2012; Mani, Mills, & Plunkett, 2012).

For the N200, there are several possible functional interpretations. First, it could be the phonological mismatch negativity effect (PMN; Connelly & Philips, 1994). The PMN, which also has a fronto-central distribution, has been elicited in both auditory as well

as picture–word priming studies (e.g., van den Brink, Brown, & Hagoort, 2001; Desroches et al., 2009, respectively). There are two possible explanations for the PMN: It either reflects a prelexical stage of word recognition based on acoustic input, or a phonological comparison of the input with an expected word form. Both explanations suggest that the PMN reflects processing of phonological form (van den Brink et al., 2001). The N200 in our study could be an instance of the PMN, since it has a similar distribution and timing. Recall that incongruous words differ from congruous words in this experiment already from the first phoneme, since all target words were paired in the training phase with words that differ in onset and vowel. Hence, the early incongruency effect could reflect the phonological mismatch between the expected (congruent) word and the presented (incongruent) word.

Sheehan et al. (2007) also put forward the N300 as a possible origin of the early semantic congruency effect. The N300, with an anterior distribution, has been reported in picture–picture priming studies (e.g., Barrett & Rugg, 1990). Its timing is 100 ms later than the N200, and typically observed in the absence of auditory words, making the N300 an unlikely source for our early semantic congruity effect.

The presence of the N200 effect implies that infants must have actively predicted the phonological word form. In other words, once infants see a novel exemplar of a trained visual category, they generate internally a label for it, even before the label is presented one second later. It is likely that this is the result of the training phase prior to test, the repetitive design of which would further encourage infants to build up expectancies of what is coming next. There is strong evidence from other ERP research in adults that people use prediction to recognize words where appropriate (e.g., van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005).

Our second semantic congruity effect, found from 400 to 600 ms is, we believe, the classical N400 effect. Whereas the earlier effect reflects a violation of the anticipated phonological word form, the later effect reflects difficulty of integrating the contrasted word within the present picture context. In other words, the N400 is sensitive to the meaning of the word. The standard time window for studying the N400 in adults is 300–500 ms, so our finding in the 400–600 ms window indicates a small delay, of around 100 ms, for infants. Moreover, the N400 does not end at 600 ms; at parietal electrodes the effect seems to have a later latency, and at other (mainly right) electrodes it extends to 800 ms. We chose the 400–600 ms range since it is the same time window studied at various ages by Mills and colleagues (e.g., Mani et al., 2012; Sheehan et al., 2007). In adults, N400 congruity effects are usually characterized by a posterior distribution. It is on these electrodes that we observe a link with infants' present vocabulary size: the larger the negativity, the more words they comprehend. This is in line with other studies showing a relation between the size of the N400 and present or subsequent vocabulary size (Friedrich & Friederici, 2006, 2010; Torkildsen et al., 2008). Note that we only observed this link for infants' present overall vocabulary size, but not for infants' specific word knowledge prior to test. This suggests that the size of the N400 is not driven by previous semantic knowledge, but rather reflects infants' ability to learn the meaning of words.

The N400 has not always been elicited with the picture–word congruency paradigm in 9-month-olds (Parise & Csibra, 2012) or even in 12-month-olds (Friedrich & Friederici, 2005a). Hence, the N400 provides an index of the rather restricted conditions under which infant semantic processing can be observed. A principal difference between our study and these two predecessor studies is the familiarization phase we implemented, in which infants gained experience both with the voice as well as with the specific word-to-world relationships. Our results do not allow us to establish what exactly boosts subsequent semantic processing:

familiarization of the voice, of the picture–word pairings, or a combination of the two. Previous research has shown that infants learn novel names for novel objects better when they first have the opportunity to see familiar objects paired with their basic level names (Fennell & Waxman, 2010; Namy & Waxman, 2000), presumably because this helps infants to better understand the purpose of the experiment (i.e. 'there will be pictures on the screen which will be named'). In addition, note that with a training phase our infants saw three times as many congruous as incongruous trials, thus potentially increasing the saliency of incongruity. Not only is it likely that the training phase boosted the word–object associations for each word, but it may thus also have strengthened picture–word priming effects in general. Friedrich and Friederici's (2004, 2005a) 12-month-olds saw three times as many incongruous trials than congruous trials (one condition of congruous trials, and three incongruous conditions: mismatching word; nonsense word; phonologically impossible word); such a ratio may have equivalently weakened the priming effects for these word–object associations that are in any case not strong. The reversed ratio in the present study could have emphasized the referential relationship between words and objects. In short the present study shows that adding a familiarization phase boosts the presence of the N400.

Recall that with the picture–word paradigm, Friedrich and Friederici also observed a larger negativity for congruous compared to incongruous words in the N200–500 range instead, which they termed the 'phonological-lexical priming effect' (Friedrich & Friederici, 2005a). This effect resembles the auditory word familiarity effect (N200–500) both in distribution, polarity and latency, consequently, it was interpreted as an index of increased word accessibility, in the context of congruous pictures. Although it is possible to view the N400 in our study as an instance of this phonological-lexical priming effect, there are several reasons why we believe the two to be distinct. First, they differ not only in polarity but also in pattern: whereas the phonological negativity effect is one large frontal effect starting 300 ms from word onset and lasting up to 600 ms, the semantic congruity effects in this study show a biphasic pattern, with one earlier short effect lasting for 100 ms (200–300 ms) followed by another longer-lasting effect (400–600 ms). Similar biphasic congruity effects have recently been observed in 14-month-olds who heard single presentations of pseudowords or mispronunciations versus correct pronunciations paired with a semantic category (Mani et al., 2012). Further, Mills and colleagues also did not report a negativity for congruous words in their studies (Mills et al., 2005).

Our results from the test phase are more comparable to the N400-like response observed in American-English speaking 14–17-month-olds as well as in adults (Mills et al., 2005). The similarities in timing and in distribution are further evidence that infants and adults use similar neural mechanisms for lexical-semantic processing. Together, the observed N200 and N400 imply that infants as young as 9 months are already capable of understanding the meaning of early words.

The present findings provide an important insight into how infants build up their first word-to-world pairings. The word learning situation was highly constrained: On each trial, infants saw one picture denoting one object and heard one spoken word produced in isolation. Further, all word tokens were produced by one speaker only (though the speaker was not present during the experiment, where words were presented via loudspeakers). In this situation, as the results showed, 9-month-olds learned the word-to-world mappings well enough to subsequently generalize across novel visual exemplars and across novel spoken words, and to accurately link these words to their referents.

There are several accounts of how infants initially learn the meaning of words (for an overview see e.g., Ambridge & Lieven,

2011). The social-pragmatic approach holds that children first need to understand the communicative intention of the speaker and have shared joint attention before they can either learn the meaning of a label or extend this label to novel exemplars (i.e., Bloom, 1993, 2000; Clark, 1993; Meltzoff, Kuhl, Movellan, & Sejnowski, 2009). Here, however, the speaker was not present, making it hard for infants to understand the speaker's communicative intentions. The 'Emergentist Coalition Model' (Hollich et al., 2000) in contrast assumes that although social and attentional cues are important for word learning, they are not mandatory, because infants are also guided by three lexical principles: reference, i.e., the knowledge that words symbolize concepts; extendibility, i.e., the knowledge that words map to more items than only the original referent; and object scope, i.e., the knowledge that words map to whole objects.

The present findings that 9-month-olds can generalize to novel exemplars provide support for the latter model. On this account, these infants already have the principles of reference as well as of extendibility. Note, however, that our findings are also compatible with association-based accounts of early word learning. These assume that infants use domain-general mechanisms to learn the meaning of words and extend them to novel exemplars by keeping track of cross-situational regularities, again without requiring infants to understand the communicative intentions (Gliozzi, Mayor, Hu, & Plunkett, 2009; Sloutsky & Robinson, 2008; Smith & Yu, 2008). Nevertheless, whether infants learn the meanings of words through increased associations between words and their referents, or through lexical principles such as that of extendibility, the present study shows that infants as young as 9 months already can perform word-to-world mappings in an adult-like manner, even for exemplars they have never seen before.

5. Conclusions

In the current study we investigated ERP responses of 9-month-olds on basic level picture–word pairings. The present research extends the literature in three substantial and significant ways. For *visual categorization*, we observed that the Nc component, associated with visual attention and recognition, is attenuated with repetition. It is only here that we observed an effect of type–token ratio of pictures: The Nc was further decreased when the picture token stays constant. This suggests that the Nc reflects here attention or saliency. For *word recognition*, the word familiarity effect became more negative with repetition, but there was no influence of picture token context, suggesting that infants have identified the concept of each picture before the word was presented (i.e., within a second). For *word-to-world mappings*, infants showed different ERP responses for words that did or did not align with the picture context, which implies that infants were able to map words to novel exemplars as young as 9 months. We observed two effects, an early N200 and an N400. The N200 implies that infants actively predicted the upcoming word form. In other words, when a novel picture of a trained category was presented, infants internally generated the phonological word form before the actual word was presented. The N400 reflects their difficulty of integrating the mismatched word with the supporting picture. Together, our results provide electrophysiological evidence of early word learning.

Acknowledgments

References

Ambridge, B., & Lieven, E. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, 14, 201–212.

Bates, E., Dale, P. S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In: P. Fletcher, & B. MacWhinney (Eds.), *Handbook of child language* (pp. 96–151). Oxford: Basil Blackwell.

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 of the United States of America*, 109, 3253–3258.

Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression. Cambridge, MA.* Cambridge University Press.

Bloom, L. (2000). Pushing the limits on theories of word learning. Commentary on Hollich, G., Hirsh-Pasek, K., & Golinkoff, R., Breaking the word learning barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65(3, serial no.262), 124–135.

Carey, S. (1978). The child as a word learner. In: M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge, UK: Cambridge University Press.

Connolly, J. F., & Philips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6, 256–266.

Courchesne, E., Ganz, L., & Norcia, A. M. (1981). Event-related brain potentials to human faces in infants. *Child Development*, 52, 804–811.

de Haan, M. (2007). *Infant EEG and event-related potentials*. New York: Psychology Press.

de Haan, M., & Nelson, C. A. (1997). Recognition of the mother's face by six-month-old infants: A neurobehavioral study. *Cognitive, Affective & Behavioral Neuroscience*, 2, 1–18.

Desroches, A., Newman, R., & Joanisse, M. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience*, 21, 1893–1906.

Duta, M. D., Styles, S. J., & Plunkett, K. (2012). ERP correlates of unexpected word forms in a picture–word study of infants and adults. *Developmental Cognitive Neuroscience*, 2, 223–234.

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infants' use of phonetic detail in word learning. *Child Development*, 81, 1376–1383.

Fenson, L., Dale, P., Reznick, J., Bates, E., Hartung, J., Pethick, S., et al. (1993). *MacArthur communicative development inventories*. San Diego, CA: Singular.

Fenson, L., Dale, P. S., Reznick, S. J., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for research in Child Development*, 59(5, serial no. 242).

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic congruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16, 1465–1477.

Friedrich, M., & Friederici, A. D. (2005a). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience*, 17, 1785–1802.

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

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, 43, 1–12.

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

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

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: Fast encoding—Weak retention. *Journal of Cognitive Neuroscience*, 23, 3228–3240.

Gliozzi, V., Mayor, J., Hu, J. F., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science*, 33, 709–738.

Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Baby wordsmith. From associationist to social sophisticate. *Current Directions in Psychological Science*, 15, 30–33.

Goyet, L., de Schonen, S., & Nazzi, T. (2010). Words and syllables in fluent speech segmentation by French-learning infants: An ERP study. *Brain Research*, 1332, 75–89.

Grossmann, T., Gliga, T., Johnson, M. H., & Mareschal, D. (2009). The neural basis of perceptual category learning in human infants. *Journal of Cognitive Neuroscience*, 21, 2276–2286.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the word learning barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65(3, serial no. 262), 65.

Jeschonek, S., Marinovic, V., Hoehl, S., Elsner, B., & Pauen, S. (2010). Do animals and furniture items elicit different brain responses in human infants? *Brain & Development*, *32*, 863–871.

Junge, C. M. M., Hagoort, P., Kooijman, V., & Cutler, A. (2010). Brain potentials for word segmentation at 7 months predict later language development. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *BUCLD 34: Proceedings of the 34th annual Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Junge, C. M. M., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. http://dx.doi.org/10.1111/j.1467-7687.2012.1144.x.

Kooijman, V., Hagoort, P., & Cutler, A. (2005). Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, *24*, 109–116.

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

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

Mani, N., Mills, D., & Plunkett, K. (2012). Vowels in early words: An event-related potential. *Developmental Science1*, *5*, 2–11.

Meltzoff, A. N., Kuhl, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science*, *325*, 284–288.

Mills, D., Coffey-Corina, S. A., & Neville, H. J. (1993). Language acquisition and cerebral specialization in 20-month-old infants. *Journal of Cognitive Neuroscience*, *5*, 317–334.

Mills, D., Coffey-Carina, S. A., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology*, *13*, 397–445.

Mills, D. L., Conboy, B., & Paton, C. (2005). Do changes in brain organization reflect shifts in symbolic functioning?. In: L. Namy (Ed.), *Symbol use and symbolic representation* (pp. 123–153). Mahwah, NJ: Lawrence Erlbaum Associates.

Mills, D., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of language and experience. *Cognitive Development*, *20*, 19–31.

Namy, L. L., & Waxman, S. R. (2000). Naming and exclaiming: Infants' sensitivity to naming contexts. *Journal of Cognition and Development*, *1*, 405–428.

Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt: Two modes of word acquisition? *Developmental Science*, *6*, 136–142.

Nelson, C. A. (1994). Neural correlates of recognition memory in the first postnatal year of life. In: G. Dawson, & K. Fischer (Eds.), *Human behaviour and the developing brain* (pp. 296–313). New York: Guilford Press.

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

Quinn, P. C., Doran, M. W., Reiss, J. E., & Hoffmann, J. E. (2010). Neural markers of subordinate-level categorization in 6-to 7-month-old infants. *Developmental Science*, *13*, 499–507.

Quinn, P. C., Westerlund, A., & Nelson, C. A. (2006). Neural markers of categorization in 6-month-old infants. *Psychological Science*, *17*, 59–66.

Reznick, J., & Kagan, J. (1983). Category detection in infancy. *Advances in Infancy Research*, *2*, 79–111.

Schafer, G. (2005). Infants can learn decontextualized words before their first birthday. *Child Development*, *76*, 87–96.

Sheehan, E. A., Namy, L. L., & Mills, D. L. (2007). Developmental changes in neural activity to familiar words and gestures. *Brain and Language*, *101*, 246–259.

Sheehan, E. A., & Mills, D. L. (2008). The effects of early word learning on brain development. In: A. D. Friederici, & G. Thierry (Eds.), *Early language development: Bridging brain and behaviour* (pp. 161–190). Amsterdam: John Benjamins.

Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science*, *32*, 342–365.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568.

Stets, M., & Reid, V. M. (2010). Infant ERP amplitudes change over the course of an experimental session: Implications for cognitive processes and methodology. *Brain and Development*, *33*, 558–568.

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175.

Tincoff, R. & Jusczyk, P. W. (2000). Do six-month-olds link sound patterns of common nouns to new exemplars? Poster presented at *the biennial meeting of the International Society for Infant Studies*, Brighton, UK.

Thierry, G., Vihman, M., & Roberts, M. (2003). Familiar words capture the attention of 11-month-olds in less than 250 ms. *Neuroreport*, *14*, 2307–2310.

Torkildsen, J. v. K., Friis Hansen, H., Svangstu, J. M., Smith, L., Simonsen, H. G., Moen, I., et al. (2008). Brain dynamics of word familiarization in 20-month-olds: Effects of productive vocabulary size. *Brain and Language*, *108I*, 73–88.

Torkildsen, J. v. K., Svangstu, J. M., Friis Hansen, H., Smith, L., Simonsen, H. G., Moen, I., et al. (2009). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience*, *20*, 1266–1282.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence form ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*, 443–467.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition, N200 versus N400 effects. *Journal of Cognitive Neuroscience*, *13*, 967–985.

Waxman, S., & Lidz, J. L., 2002. Early word learning. In D. Kuhn & R. Siedler (Eds.), *Handbook of child psychology* (6th ed.) (pp. 299–335). Hoboken, NJ: Wiley.

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

Zink, I., & Lejaegere, M. (2002). *N-CDIs: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al.* Leuven, Belgium/Leusden, Netherlands: Acco.