Electrophysiological Evidence for the Involvement of the Approximate Number System in Preschoolers' Processing of Spoken Number Words

Michal Pinhas, Sarah E. Donohue, Marty G. Woldorff,
and Elizabeth M. Brannon

Abstract

■ Little is known about the neural underpinnings of number word comprehension in young children. Here we investigated the neural processing of these words during the crucial developmental window in which children learn their meanings and asked whether such processing relies on the Approximate Number System. ERPs were recorded as 3- to 5-year-old children heard the words one, two, three, or six while looking at pictures of 1, 2, 3, or 6 objects. The auditory number word was incongruent with the number of visual objects on half the trials and congruent on the other half. Children's number word comprehension predicted their ERP incongruency effects. Specifically, children with the least number word knowledge did not show any ERP incongruency effects, whereas those with intermediate and high number word knowledge showed an enhanced, negative polarity incongruency response (N_{inc}) over centroparietal sites from 200 to 500 msec after the number word onset. This negativity was followed by an enhanced, positive polarity incongruency effect (P_{inc}) that emerged bilaterally over parietal sites at about 700 msec. Moreover, children with the most number word knowledge showed ratio dependence in the P_{inc} (larger for greater compared with smaller numerical mismatches), a hallmark of the Approximate Number System. Importantly, a similar modulation of the P_{inc} from 700 to 800 msec was found in children with intermediate number word knowledge. These results provide the first neural correlates of spoken number word comprehension in preschoolers and are consistent with the view that children map number words onto approximate number representations before they fully master the verbal count list. ■

INTRODUCTION

Learning the meaning of number words is a gradual process that goes well beyond the age at which children first memorize the count list (Wynn, 1990, 1992). Children at the age of 2 are typically able to recite the count list, but do so with little understanding of how it relates to the number of elements in a set. Accordingly, young children exhibit only partial knowledge about the cardinal meaning of the counted number words (e.g., understanding just *one*, just *one* and *two*, and so on) and are considered "subset-knowers." It is only around the age of 4 that children understand that the last number reached when counting a set establishes its cardinality and are thus considered "cardinal principle knowers" (CP-knowers; e.g., Condry & Spelke, 2008; Sarnecka & Carey, 2008; Le Corre & Carey, 2007; Le Corre, Van de Walle, Brannon, & Carey, 2006; Wynn, 1990, 1992).

Before children learn the count sequence, they already show robust evidence of representing number approximately without language. In fact, the Approximate Number System (ANS) is seen in a wide array of species and emerges early in human infancy (Feigenson, Dehaene, & Spelke, 2004). A hallmark of the ANS is ratio dependence, that is, discrimination is dependent on the ratio between numerosities and not their absolute values. Thus, empirically ratio dependence is considered a litmus test for the ANS. Despite the existing behavioral and neural evidence that the ANS is employed in number processing in infancy (e.g., Libertus & Brannon, 2009; Berger, Tzur, & Posner, 2006; Brannon, Abbott, & Lutz, 2004; Feigenson et al., 2004; Xu & Spelke, 2000), there is much controversy over the nature of the nonverbal representations that serve in childhood as the basis for learning the meaning of number words (e.g., Wagner & Johnson, 2011; Carey, 2009; Sarnecka & Lee, 2009; Le Corre & Carey, 2007; Slaughter, Kamppi, & Paynter, 2006; Rousselle, Palmers, & Noel, 2004; Huntley-Fenner & Cannon, 2000). According to one view, the ANS serves as the foundation upon which arbitrary number words are mapped (e.g., Gelman & Gallistel, 2004; Pica, Lemer, Izard, & Dehaene, 2004; Gallistel & Gelman, 1990). A second view holds that children map number words onto enriched parallel individuation representations of up to three or four single items. Accordingly, this proposed object file system encodes number implicitly, such that there are no set level numerical

Duke University

Journal of Cognitive Neuroscience 26:9, pp. 1891–1904
doi:10.1162/jocn_a_00631

representations, and instead each object in a set is represented individually (e.g., Carey, 2009; Sarnecka & Lee, 2009; Le Corre & Carey, 2007). Furthermore, although there has been considered behavioral work pursuing resolution of this issue, no previous studies have investigated the neural underpinnings of the processing of spoken number words in young children and whether there is neural evidence that such representations are mapped onto the ANS.

Approximate number representations are also considered by many researchers to be abstract and supramodal (but see Cohen Kadosh & Walsh, 2009). That is, performance depends on number magnitude, regardless of the specific format or modality (e.g., symbolic number: spoken or written number words, or nonsymbolic number: visual arrays, sequence of tones) in which values are presented (e.g., Feigenson et al., 2004; Dehaene, Dehaene-Lambertz, & Cohen, 1998). In fact, infants show ratio dependence discrimination of (nonsymbolic) numerosity in both visually and auditorily presented arrays (e.g., Feigenson et al., 2004; Lipton & Spelke, 2003). Furthermore, data from children and adults demonstrate number-related effects across modalities and across both nonsymbolic and symbolic notations (e.g., Libertus, Woldorff, & Brannon, 2007; Temple & Posner, 1998; for a review, see Cohen Kadosh & Walsh, 2009).

Previous ERP studies examining the development of word comprehension have used semantic incongruency paradigms with multimodal stimuli to study how the child's brain processes words in meaningful contexts. The basic reasoning of this approach has been that detection of semantic incongruencies would depend on the child's word comprehension level. In these studies, a picture of an object is paired with an auditory word that is either semantically congruent or incongruent with the visual stimulus. The main two ERP effects that have been described for such incongruency paradigms are the N400 and LPC (for "late positive complex"). The N400 is a language-related negative polarity potential that peaks about 400 msec poststimulus over centroparietal sites. It is a hallmark electrical brain response that is elicited in response to wide variations of meaningful stimuli and is thought to reflect semantic integration into context. The N400 is larger in amplitude for semantically incongruent than congruent trials, thereby marking the detection of a contextual mismatch, with its size generally reflecting the degree of the mismatch (for a review, see Kutas & Federmeier, 2011). The elicitation of an N400 has been reported in the picture–word paradigm in infants (although usually much delayed than the typical latency in adults), older children, and adults (e.g., Henderson, Baseler, Clarke, Watson, & Snowling, 2011; McCleery et al., 2010; Friedrich & Friederici, 2004; Byrne et al., 1999).

The LPC is also elicited in both children and adults in incongruency paradigms, typically occurring at ~600–900 msec poststimulus over midline and posterior sites in response to meaningful visual and auditory stimuli (e.g., Sitnikova, Kuperberg, & Holcomb, 2003; Juottonen, Revonsuo, & Lang, 1996; Friedman, Sutton, Putnam, Brown, & Erlenmeyer-Kimling, 1988; Licht, Kok, Bakker, & Bouma, 1988; Kok & Rooijakkers, 1985). Such late positive waves are elicited under different task conditions and show variability in peak latencies and scalp distributions, suggesting that they are unlikely to reflect a single cognitive process. Given that the LPC can be evoked by semantic violations, with increased amplitude to semantic mismatches, it is also thought to reflect the later processing of stimulus meaning (Munte, Heinze, Matzke, Wieringa, & Johannes, 1998; Juottonen et al., 1996).

Incongruency-related ERP effects are also elicited by paradigms used to study stimulus conflict processing, such as in the classic Stroop task (Stroop, 1935), where participants are presented with color words and respond to the color of the font, which can be congruent or incongruent with the meaning of the word. Congruent trials are responded to faster than incongruent trials, and the RT difference between these trials is called the Stroop effect. Two ERP incongruency effects are typically elicited in this task: a centrally distributed negative polarity wave peaking ~450 msec (N450), followed by a later (500–800 msec) posterior sustained positivity (SP; e.g., Appelbaum, Meyerhoff, & Woldorff, 2009; Larson, Kaufman, & Perlstein, 2009; West, 2003; Liotti, Woldorff, Perez, & Mayberg, 2000). Similar components have been identified in an auditory version of the Stroop task, supporting supramodal conflict detection mechanisms (Donohue, Liotti, Perez, & Woldorff, 2012). Furthermore, both of these ERP effects have been reported in adults across various response modes, including where no verbal or manual responses were required (Donohue et al., 2012; Liotti et al., 2000). The N450 or the N_{inc} (for negative polarity incongruency effect; Donohue et al., 2012) has been hypothesized to be related to conflict-processing activity in the ACC (Szücs & Soltesz, 2012; West, 2003; Liotti et al., 2000) as well as to resemble the N400 by indexing semantic conflict at the stimulus level (Rebai, Bernard, & Lannou, 1997). The late SP component has been proposed to reflect additional semantic processing by posterior word processing areas (Liotti et al., 2000), control processes related to trial completion (Larson et al., 2009), and represent activity of a supramodal attentional control system that facilitates and suppresses task-relevant and -irrelevant information (Donohue et al., 2012), respectively.

Similar incongruency ERP effects marked by a negativity (N450) followed by somewhat shorter-duration positivity have been reported in adults using a numerical version of the Stroop task using Arabic numerals (e.g., Szücs & Soltesz, 2007, 2012; Szücs, Soltesz, & White, 2009). ERP interference effects to Arabic numerals have also been observed in school-aged children, starting from the first grade (e.g., Soltesz, White, & Szücs, 2011; Szücs, Soltesz, Jarmi, & Csepe, 2007). The numerical Stroop task and other numerical processing tasks with children and adults have been shown to elicit ERPs that are modulated by numerical distance or ratio, usually distributed over parietal sites. Such effects were reported at both early and/or later latencies (e.g., Soltesz et al., 2011; Hsu & Szücs, 2010; Paulsen, Woldorff, & Brannon, 2010; Hyde & Spelke, 2009; Libertus et al., 2007; Szücs & Soltesz, 2007; Szücs et al., 2007; Temple & Posner, 1998; Dehaene, 1996), providing neural evidence for the involvement of the ANS in both nonsymbolic and symbolic numerical forms. To our knowledge, only one prior ERP study in adults specifically examined the neural response to spoken number words, which are the first acquired form of symbolic number (Szücs & Csepe, 2005). That study reported ERP distance effects over frontal and parietal sites at ~200 msec poststimulus. Thus, numerical distance-related ERPs have been found with both visual and auditory stimuli in paradigms that require participants to make numerical judgments as well as in more passive-viewing designs.

The goals of this study were twofold. The first goal was to examine whether the brains of young children show incongruency-related ERP effects to spoken number words when they did not match a visual object scene. The second goal was to explore the brain correlates that are associated with children's number word knowledge and whether these correlates are indicative of mapping number words onto ANS representations as would be evidenced by numerical ratio ERP effects. Whereas we would not predict that a child who did not know the words chinchilla and avocado would show an incongruency effect for the juxtaposition of one of these words with one of the objects the words represent, number is a different case. As described earlier, children know the number words well before they are able to precisely construct sets that represent the values of the words. Children may have implicit knowledge of the meaning of the number words that could be detected by ERP incongruency effects. If those incongruency effects also showed ratio effects, this could provide evidence that approximate representations of number are being mapped onto number words before children show master of the cardinal principle.

To this end, we implemented a version of the ERP picture–word paradigm (e.g., Henderson et al., 2011; McCleery et al., 2010; Friedrich & Friederici, 2004; Byrne et al., 1999) to test children's neural response to spoken number words that were either congruent or incongruent with the number of visible objects. Children's number word comprehension level was measured by the Give-a-Number procedure (Wynn, 1990, 1992), and we also assessed their general verbal ability (Developmental Vocabulary Assessment for Parents [DVAP]; Libertus, Odic, Feigenson, & Halberda, 2013). To target children who showed a wide range in mastery of the verbal count list, we tested younger children at the preschool age range (from 3 to 5 years) than previously studied on neural processing of symbolic number.

Given the age of the children and the fact that many children were expected to show behavioral evidence of understanding just a few numbers words (e.g., only "one" and "two"), we limited the stimuli to the first number words that a typical child acquires ("one," "two," and "three") and included "six" to increase the possible ratios for the later testing of numerical ratio effects. It is important to note that, although three of these four words represent small values within the subitizing range, prior studies across development and across species have shown that small numerical values can elicit ratio-dependent responses, consistent with the employment of the ANS (e.g., Cordes & Brannon, 2009; Cordes, Gelman, Gallistel, & Whalen, 2001; Brannon & Terrace, 1998).

On the basis of previous findings, we hypothesized that children's understanding of number words would modulate their integration of semantic information into the numerically incongruent versus congruent context, which would be reflected by a negative polarity incongruency ERP effect and a late positive incongruency effect, both known to reflect the detection of semantic incongruency. We predicted that CP-knowers, who understood the meaning of all number words presented, would show larger ERP incongruency effects compared with children who exhibited only partial understanding. Furthermore, because the semantic processing of incongruent trials hinges on the comparison of the number word with the number of visual objects, we expected parietally distributed multimodal numerical ratio effect instantiated by a greater incongruency for larger compared with smaller ratio trials (e.g., 2 objects paired with "six" vs. 3 objects paired with "six," respectively; e.g., Soltesz et al., 2011; Hsu & Szücs, 2010; Paulsen et al., 2010; Hyde & Spelke, 2009; Libertus et al., 2007; Szücs & Soltesz, 2007; Szücs et al., 2007; Szücs & Csepe, 2005; Temple & Posner, 1998; Dehaene, 1996). Such modulation by numerical ratio would implicate the use of the ANS in young children's processing of spoken number words. A secondary question was whether subset-knowers would show any evidence of the same numerical ratio ERP effects as CP-knowers. Such evidence would suggest that number words are mapped onto ANS representations during the formative period where children are learning the meaning of number words rather than after this process is fully established.

METHODS

Participants

One hundred fifteen children (3.11–5.57 years, mean = 3.98, 59 girls, 6 bilingual) from the greater Research Triangle, NC, community were included in the final analysis. Data from 35 additional children were discarded (see EEG Recording and Analysis section for details). All parents provided written informed consent for their children, using a protocol approved by the Duke University Institutional Review Board. Children received a small toy as a reward for participation. Parents were compensated for their time (~$45).

Stimuli and Apparatus

EEG Picture/Number Word Task

The visual stimuli consisted of color photographs of balls, chairs, dogs, and shoes presented as arrays of one (~3.6° high and 3.1° wide), two (~3.6° high and 6.2 wide), three (~3.6° high and 9.3 wide), or six (~7.1° high and 9.3° wide) objects on a white background. To increase interest, four different exemplars of each category were used. Overall, there were 64 different visual images (4 exemplars × 4 categories × 4 numerical values). A black fixation cross (0.6° × 0.6°) was present at the center of the screen at all times.

The auditory stimuli were two-word phrases spoken by a recorded female voice. Each phrase began with the word "look," followed by "one"/"two"/"three"/"six" (mean word length = 409 msec; 10-msec rise and fall periods). The auditory stimuli were presented centrally through two speakers with an intensity of ~55 dB SPL. The presentation of the stimuli was controlled by Presentation software package (Neurobehavioral Systems, Albany, CA).

Each block consisted of 96 trials, with numerically congruent and incongruent trials presented with equal frequency. On congruent trials, the spoken number word matched the number of items in the picture (48 trials = 4 number words × 4 item types × 3 repetitions), whereas on incongruent trials it was inconsistent with the number of visual objects (48 trials = 4 number words × 3 items number × 4 item types). The exemplar for each object category was randomly selected (with replacement) on each trial. Thus, for a visual image of two objects, the accompanying auditory stimulus would be "Look, two!" on congruent trials and "Look, one/three/six!" on incongruent trials. The 48 incongruent trials consisted of 12 trial types and 6 unique numerical contrasts (e.g., collapsing trials with "one" paired with 6 objects and trials with "six" paired with 1 object). Further collapsing by ratio allowed us to consider pairs with a 1:2 (pairs 1:2 and 3:6), 1:3 (pairs 1:3 and 2:6), 1:6 (pair 1:6), or 2:3 (pair 2:3) ratio between the number word and the number of objects.

Procedure

Children were first tested in the behavioral "Give-a-Number" task (Wynn, 1990, 1992), followed by the EEG picture–number word study.[1]

Verbal Ability

Parents completed the DVAP (Libertus et al., 2013) during their child's testing session. The DVAP lists the first 212 words of Form A of the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007). Parents indicated the listed words they had previously heard their child say. The total number of words indicated represents the child's DVAP score with higher scores reflecting larger vocabulary sizes. Libertus et al. (2013) found that preschoolers' performance on the PPVT-4 was highly correlated with their DVAP scores (see also Libertus, Feigenson, & Halberda, 2011); thus, we used the DVAP to shorten children's testing time.

Behavioral Assessment: Give-a-Number

The experimenter asked the child to give a puppet different numbers of small toys as part of a game. First, the experimenter placed the puppet, a plate, and pile of fish-shaped erasers on a table in front of the child and stated "In this game you will give fish to Elmo, like this," while placing two fish on the plate and sliding it over to the puppet. The experimenter then placed the fish back in the pile and said "Can you give Elmo *one* fish?" After the child presented the plate to the puppet the experimenter asked "Is that *one*?" If the child said "Yes," the experimenter said "Thank you!" and placed the toys back in the pile. If the child said "No," the experimenter repeated the trial again from the start. No feedback was provided. Children were allowed to fix or change their response and there was no time limit.

Children were asked for one fish on the first trial and for three fish on the second. If the child succeeded on both trials, then five fish were asked for; if the child failed, then two fish were asked for. Subsequent requests depended on the child's previous responses. If she or he succeeded at a given number (N), then $N + 1$ was requested at the next trial, if not then $N - 1$ was requested. Requests continued until the child had at least two successes at a given N and at least two failures at $N + 1$. The highest number requested was 6.

Children were categorized by knower level based on the highest number word they knew. Children who reliably generated sets of five or fewer items were considered subset-knowers and categorized as one-, two-, three-, four-, or five-knowers, depending on the highest number they successfully gave. For example, children who successfully constructed sets of one and two toys, but not three toys, were called two-knowers. Children who successfully constructed sets of six toys were categorized as CP-knowers (Wynn, 1990, 1992). For the purposes of the statistical analyses of the EEG task, subset-knowers were divided into two groups: one- and two-knowers and three- to five-knowers, and their performance were compared with that of CP-knowers.

EEG Picture/Number Word Task

Participants were seated in an electrically shielded booth, approximately 74 cm from the center of a 20-in. monitor. An experimenter sat in another chair, to the right of the child, and monitored the child's behavior. Children were instructed to attend to the screen and avoid movements. No overt responses were required. On each trial, a picture with one, two, three, or six objects appeared on the center of the computer screen for 2500 msec. The spoken word "look" was played at the onset of the picture to draw the child's attention to the image. Nine hundred milliseconds after trial onset, while the picture was still present on the screen, a spoken number word was played. The 900-msec delay between the onset of the auditory and visual stimuli allowed for the evoked responses to the visual stimuli to return to a baseline level before the onset of the stimulus of interest (i.e., the spoken number word) and, moreover, the type of this latter stimulus (e.g., congruent vs. incongruent) was randomized. The intertrial interval varied between 1.1 and 1.3 sec and displayed a white screen with central black fixation cross (Figure 1). Note that children were not allowed to move their hands to count the objects presented in the picture, and we also monitored their mouth movements to insure that they were not counting verbally. Each experimental block started with the presentation of a congruent trial that was excluded from analysis. The remaining 96 trials within the block were composed of 50% congruent and 50% incongruent trials. The frequency of each number word and each numerical array of objects was equally distributed in each block. An animated movie was presented on the computer monitor for 40 sec after every 24 trials with the goal of increasing children's interest in looking at the screen. Each block lasted approximately 8 min. During a short break between blocks, children were offered small rewards (e.g., snacks, stickers) and given feedback about her or his behavior (e.g., attentiveness, ability to sit still, etc.). EEG recordings were stopped during the blocks if the experimenter noticed that the child lost interest in the task, was moving a lot, or was not looking at the monitor. In those cases, a short break was taken, and the experiment resumed when the child was ready to continue. Fifty-three children completed four blocks (384 trials total). The remaining 62 children completed three blocks (288 trials total), because of loss of interest, impatience, fatigue, and so on.

EEG Recording and Analysis

Ongoing EEG was recorded from 32 channels mounted in a customized, elastic electrode cap (Duke32 Waveguard cap layout, Advanced Neuro Technology, the Netherlands). Channels were equally spaced across the cap and covered the whole head from above the eyebrows to below the inion. The EOG was monitored with an electrode below the left eye (vertical EOG) and two other electrodes on the left and right canthi (horizontal EOG). Recordings were referenced online to the average of all channels, low-pass filtered at 138 Hz, and digitized with a 512-Hz

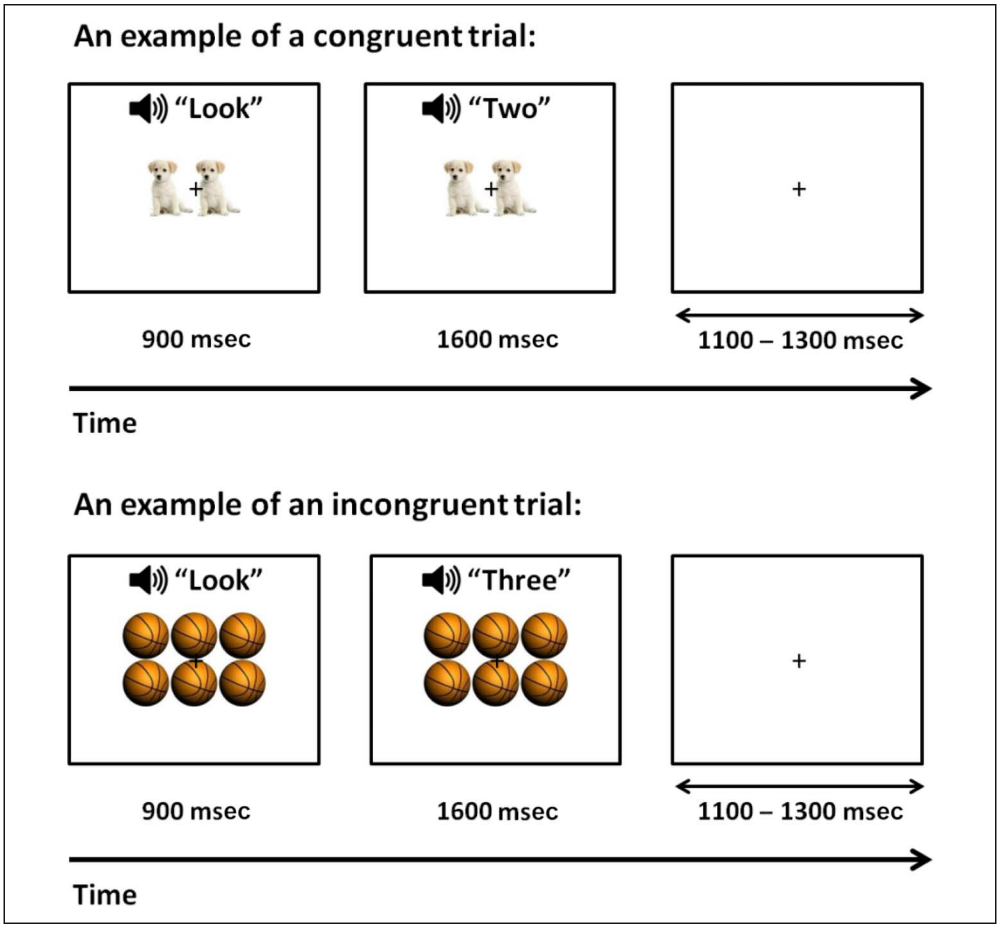

Figure 1. Illustration of trial structure, duration, and stimuli in the EEG picture/number word task. In congruent trials (top), the spoken number word matched the number of items in the picture, whereas in incongruent trials (bottom) the number word did not match the number of items in the picture. The extracted ERPs were time-locked to the onset of the spoken number word.

sampling rate. Electrode impedances were maintained below 10 kΩ for the EOG channels, below 2 kΩ for the ground and left and right mastoid channels, and below 5 kΩ for all other channels.

Offline, data were subjected to artifact detection and rejection (implemented by an automated computer algorithm after the manual setting of artifact amplitude rejection levels). Trials with extreme eye artifacts and off-scale activity (muscle artifact, movement artifact, or electrode drift) were excluded. The data were then digitally band-pass filtered between 0.016 and 30 Hz and rereferenced to the algebraic average of the left and right mastoids. Artifact-free trials for each participant were next segmented into epochs beginning from 500 msec before the onset of the spoken number word to the end of the trial, averaged into experimental conditions of interest (e.g., congruent, incongruent), and baseline corrected to the average amplitude from −200 to 0 msec before the spoken word onset. To test for the numerical ratio effect, incongruent trials were further segregated into four ratio conditions of 2:3 ("two"/"three" paired with 3/2 objects, respectively), 1:2 ("one"/"two" paired with 2/1 objects and "three"/"six" paired with 6/3 objects, respectively), 1:3 ("one"/"three" paired with 3/1 objects and "three"/"six" paired with 6/3 objects, respectively), and 1:6 ("one"/"six" paired with 6/1 objects, respectively). Data from 35 of the participants were excluded because they had fewer than 50 usable trials for either the congruent or the incongruent condition, leaving 115 participants in the full analyses. On average 103.2 (*SD* = 31.9) congruent and 103.3 (*SD* = 31.1) incongruent trials were retained for each child. The average number of 2:3, 1:2, 1:3, and 1:6 incongruent ratio trials retained for each child was 17.0 (*SD* = 5.3), 34.3 (*SD* = 10.5), 34.4 (*SD* = 10.4), and 17.6 (*SD* = 5.5), respectively. Finally, grand averages across all participants, as well as for the three different subgroups of knower level (1–2 knowers, 3–5 knowers, CP-knowers), were computed for each congruency condition. In addition, incongruent minus congruent difference waves were computed for each ratio relationship,[2] as well as collapsed across ratio conditions for each of the knower level groups. Topographical maps for each of the difference waves were also generated for each participant group and across all participants.

Because no previous study had examined ERPs in young children in response to number words, we had no precise metric upon which to rely for the statistical selection of the channels and time windows to test. Nevertheless, based on previous work using related incongruency paradigms, we expected that there would be an early centrally distributed incongruency-related negativity and a later posterior parietal positivity reflecting the neural processing of the mismatch of our audiovisual stimuli. Visual inspection of the grand-averaged waveforms across participants for incongruent and congruent trials, and their difference waves revealed two main effects of incongruency consistent with these expectations—namely, a relatively early central-parietal incongruency-related negativity, followed by a late bilateral parietal positivity. The enhanced negativity, hereafter referred to as the N_{inc} (for incongruency negativity), resembled the negative incongruency effects previously found with picture–word and word–color Stroop ERP paradigms (e.g., Donohue et al., 2012; Henderson et al., 2011; Appelbaum et al., 2009; Szücs et al., 2009; Friedrich & Friederici, 2004; West, 2003; Liotti et al., 2000; Byrne et al., 1999). The N_{inc} started ~150 msec earlier here than the N450 typically observed in adults with Stroop paradigms, possibly because of conflict processing in the auditory modality, which has been found to begin earlier (Donohue et al., 2012). Guided by these previous studies and the visual inspection of the data, we analyzed the N_{inc} from 200 to 500 msec, sampling from three adjacent, centroparietal, midline electrodes, as well as three immediately adjacent electrodes to the left and right of this midline (Figure 2).

Similarly, the longer-latency positivity incongruency effect, hereafter referred to as the P_{inc} (for incongruency positivity), resembled the late sustained positive polarity effect previously reported to follow the N400, N_{inc}, and N450 effects (e.g., Donohue et al., 2012; Appelbaum et al., 2009; Szücs & Soltesz, 2007; Sitnikova et al., 2003; West, 2003; Liotti et al., 2000; Juottonen et al., 1996; Friedman et al., 1988), although the effect seemed to start ~100 msec later here for the CP-knowers and ~200 msec later for the 3–5 knowers. Accordingly, the P_{inc} was analyzed from 700 to 1000 msec, sampling from three adjacent, more laterally located parietal electrodes on the left, and the three corresponding group of lateral parietal electrodes on the right (Figure 3).

A repeated-measures ANOVA on mean amplitude values with Knower Level as a between-participant variable (1–2 knowers, 3–5 knowers, CP-knowers) and Congruency (congruent, incongruent) and Laterality (for N_{inc}: left, midline, right; for P_{inc}: left, right) as within-participant variables was conducted for both the N_{inc} (200–500 msec) and the P_{inc} (700–1000 msec) effects. In cases Congruency interacted with Knower Level, follow-up tests evaluated the significance of the congruency effect for each knower level group as well as contrasted its strength between knower level groups. Additional tests evaluated the prediction of a numerical ratio effect in cases of significant N_{inc} and P_{inc} effects by computing linear trends on the mean amplitude of the difference waves (incongruent minus congruent) of the four ratio conditions. In cases of significant numerical ratio effects, additional comparisons were conducted on the mean amplitudes of the difference waves to test for an effect of numerical pair in the ratio conditions 1:2 (contrasting 1:2 pairs vs. 3:6 pairs) and 1:3 (contrasting 1:3 pairs vs. 2:6 pairs). These analyses tested the possibility that the absolute values (i.e., the specific ratio magnitudes) modulated the effects. Subsequent *t* tests evaluated potential effects of absolute values by resegmenting incongruent trials to small (the pairs 1:2, 1:3 and 2:3) versus large values (the pairs 1:6, 2:6 and

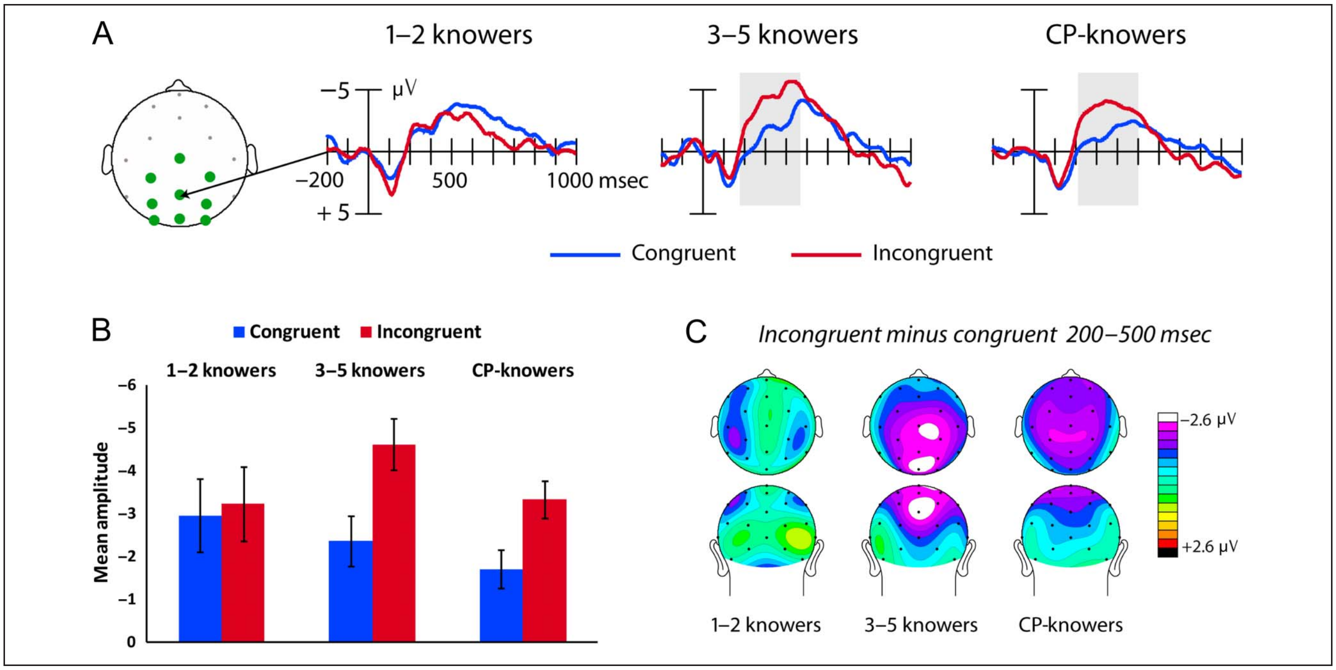

Figure 2. (A) ERP traces for congruent and incongruent conditions by knower level, time-locked to the onset of the spoken number word. An early negative wave incongruency effect (N_{inc}) emerged starting at around 200 msec. Each tick mark represents 100 msec. Sites used for testing the N_{inc} effect are highlighted. (B) Mean amplitude values (microvolts) of congruency by knower level. Error bars denote .95 confidence intervals. (C) Topographic distributions for incongruent minus congruent trials from 200 to 500 msec. The N_{inc} effect was found in both 3–5 knowers and CP-knowers and had a central distribution.

3:6). Lastly, we calculated the correlations (Spearman's rho or Pearson's *r*) between the child's knower level (six levels: from one-knower to CP-knower), individual verbal ability (DVAP scores), age, and the N_{inc} and P_{inc} effects (calculated as the mean amplitudes of incongruent minus congruent difference waves). Significance level was defined as $p < .05$ and inferred for Greenhouse–Geisser corrected *p* values in all ANOVAs.

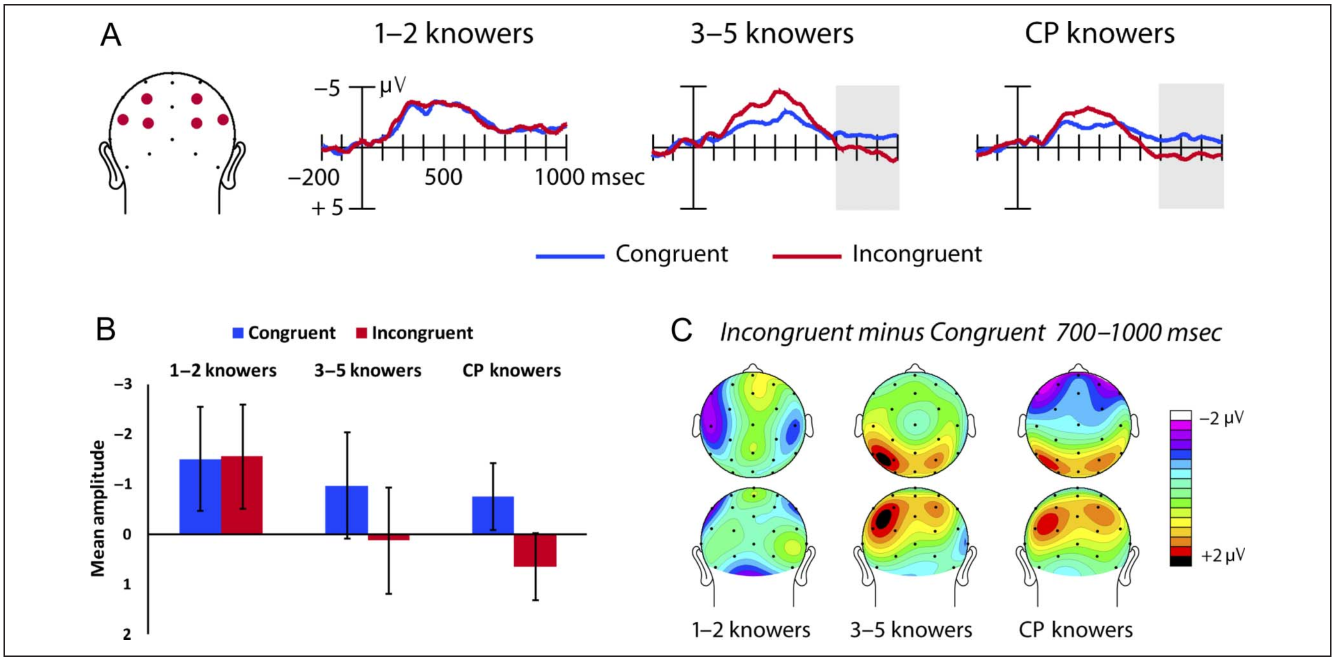

Figure 3. (A) ERP traces averaged across the six highlighted channels by knower level for congruent and incongruent conditions. Traces are time-locked to the onset of the spoken number word. A late positive wave incongruency effect (P_{inc}) emerged at around 700 msec. Each tick mark represents 100 msec. (B) Mean amplitude values (microvolts) of congruency by knower level. Error bars denote .95 confidence intervals. (C) Topographic distributions for incongruent minus congruent trials from 700 to 1000 msec. The P_{inc} effect was parietally distributed and observed in both 3–5 knowers and CP-knowers.

RESULTS

Verbal Ability

Parents reported that their children used 89.3 words on average (*SD* = 27.0). Similar averaged DVAP score in preschoolers was reported by Libertus et al. (2011). DVAP scores significantly correlated with age (Pearson's $r = .41$, $p = .01$), such that older children were evaluated as having larger vocabulary sizes than younger children.

Give-a-Number Behavioral Task

Categorization based on the knower level framework (Wynn, 1990, 1992) yielded 59 subset-knowers (7 one-knowers, 24 two-knowers, 13 three-knowers, 7 four-knowers, 8 five-knowers) and 56 CP-knowers. The subset-knowers were divided into two groups of 1–2 knowers (31 children; 3.11–4.06 years, mean = 3.37, 13 girls, 4 bilingual) and 3–5 knowers (28 children; 3.15 to 5.19 years, mean = 3.69, 17 girls), and their EEG data were compared with each other and with that of CP-knowers (56 children; 3.22–5.57 years, mean = 4.46, 29 girls, 2 bilingual). As expected, knower level correlated with age (Spearman's rho = .69, $p < .00001$), reflecting that older children knew the meaning of more number words than did younger children.

EEG Picture/Number Word Task

The data revealed two incongruency effects, an early centrally distributed negativity (N_{inc}) and a later parietally distributed positivity (P_{inc}), that were modulated by children's knower level. Note that main effects of Laterality and Knower Level (i.e., on the raw ERP waveforms) are not reported because of their lack of theoretical importance to the current questions.

Table 1. Overview of the Statistical Analysis of Absolute Value Effects

Comparison	*df*	*t*	*SD*	*p*
N_{inc}				
3–5 knowers				
Congruent vs. incongruent small values	27	−4.2	3.0	.0002
Congruent vs. incongruent large values	27	−3.7	3.0	.0009
Incongruent small values vs. large values	27	−0.4	4.2	.72
CP-knowers				
Congruent vs. incongruent small values	55	−4.1	3.4	.0001
Congruent vs. incongruent large values	55	−4.4	2.4	.00004
Incongruent small values vs. large values	55	−0.9	3.7	.35
P_{inc}				
3–5 knowers				
Congruent vs. incongruent small values	27	2.0	3.6	.06
Congruent vs. incongruent large values	27	2.4	2.9	.02
Incongruent small values vs. large values	27	0.004	3.5	.99
CP-knowers				
Congruent vs. incongruent small values	55	2.4	3.1	.02
Congruent vs. incongruent large values	55	4.1	3.3	.0002
Incongruent small values vs. large values	55	1.5	4.1	.15

N_{inc} (200–500 msec)

A significant main effect of Congruency [$F(1, 112) = 30.1$, $MSE = 10.1$, $p < .000001$, $\eta^2_p = .21$] manifested as a greater ERP negativity in the incongruent compared with the congruent condition. Importantly, Congruency interacted with Knower Level [$F(2, 112) = 4.7$, $MSE = 10.1$, $p = .01$, $\eta^2_p = .08$]. Planned comparisons revealed that the N_{inc} was significant for both 3–5 knowers [$F(1, 27) = 30.8$, $MSE = 2.3$, $p = .000007$, $\eta^2_p = .53$] and CP-knowers [$F(1, 55) = 27.8$, $MSE = 2.6$, $p = .000002$, $\eta^2_p = .34$], but not for 1–2 knowers ($F < 1$, $\eta^2_p = .007$; Figure 2). Next, we compared the strength of the N_{inc} between the groups that produced a significant effect (i.e., 3–5 knowers and CP-knowers) with the group that did not (i.e., 1–2 knowers) and found a significant difference [$F(1, 112) = 9.14$, $MSE = 10.0$, $p = .003$, $\eta^2_p = .08$]. Furthermore, there was no difference in the strength of the N_{inc} found for 3–5 knowers and CP-knowers [$F(1, 112) = 1.09$, $MSE = 10.0$, $p = .29$, $\eta^2_p = .009$]. Additional comparisons for the groups that produced a significant N_{inc} did not reveal numerical ratio effects on this component, as reflected by the linear trends not being significant within those conditions (3–5 knowers: $F < 1$, $\eta^2_p = .02$; CP-knowers: $F < 1$, $\eta^2_p = .009$). Furthermore, the segregation of incongruent trials into small versus large value trials did not provide any alternative evidence for any absolute value effects in either group. Whereas a significant N_{inc} was present when congruent trials were contrasted with incongruent small-value trials as well as with incongruent large-value trials ($ps < .001$), the two types of incongruent trials did not differ significantly from each other (see Table 1). Furthermore, no other interactions with congruency were found to be significant for the N_{inc}.

P_{inc} (700–1000 msec)

A significant main effect of Congruency manifested as an enhanced parietal positivity for incongruent compared with congruent trials [$F(1, 112) = 11.3$, $MSE = 7.0$, $p = .001$, $\eta^2_p = .09$]. Importantly, as in the N_{inc} analysis, congruency was modulated by the knower level [$F(2, 112) = 3.4$, $MSE = 7.0$, $p = .04$, $\eta^2_p = .06$]. This interaction resulted from there being a significant P_{inc} effect for both 3–5 knowers [$F(1, 27) = 6.3$, $MSE = 7.5$, $p = .02$, $\eta^2_p = .19$] and CP-knowers [$F(1, 55) = 17.6$, $MSE = 6.2$, $p = .0001$, $\eta^2_p = .24$], but not for 1–2 knowers ($F < 1$, $\eta^2_p = .0008$; Figure 3). Furthermore, although there was a significant difference in the strength of the P_{inc} found for the groups that produced a significant effect (i.e., 3–5 knowers and CP-knowers) as contrasted with the group that did not (i.e., 1–2 knowers) [$F(1, 112) = 6.33$, $MSE = 7.0$, $p = .01$, $\eta^2_p = .05$], there was no difference in the strength of the P_{inc} activations found for 3–5 knowers and for CP-knowers ($F < 1$, $\eta^2_p = .0002$). The P_{inc} manifested as a bilateral parietal positivity for the CP-knowers and seemed more left lateralized for the 3–5 knowers, although no significant Congruency × Laterality interaction was found for the latter group of participants [$F(1, 27) = 2.7$, $MSE = 2.4$, $p = .11$, $\eta^2_p = .09$]. In contrast to the N_{inc}, the P_{inc} activity did show effects of numerical ratio. Planned comparisons on the P_{inc} difference waves (incongruent minus congruent) revealed a significant linear trend for the CP-knowers [$F(1, 55) = 4.59$, $MSE = 57.2$, $p = .04$, $\eta^2_p = .08$], demonstrating a gradual increase in the P_{inc} amplitudes with the increase in the numerical mismatch. A similar pattern was observed for the 3–5 knowers but did not reach significance across the full 700–1000 msec latency window [$F(1, 27) = 3.13$, $MSE = 53.4$, $p = .09$, $\eta^2_p = .10$]. Given the theoretical importance of finding a ratio effect in subset-knowers, we next tested the significance of the same linear trend over smaller, consecutive 100 msec time windows. This analysis revealed that a significant ratio effect was indeed present in 3–5 knowers P_{inc} from 700 to 800 msec poststimulus [$F(1, 27) = 6.29$, $MSE = 52.6$, $p = .02$, $\eta^2_p = .19$], reflecting a monotonic decrease in the mean amplitudes of the P_{inc} with the decrease in numerical ratio. The ratio effect was marginally significant in the following 800–900 msec time window [$F(1, 27) = 3.84$, $MSE = 62.4$, $p = .06$, $\eta^2_p = .12$] and not significant from 900 to 1000 msec ($F < 1$, $\eta^2_p = .01$). Figure 4 depicts the modulations of the P_{inc} by numerical ratio for each knower level group. To complete the picture, we found no absolute value effects in the P_{inc} (700–1000 msec) when contrasting the mean amplitude of the difference waves calculated for the unique pairs constituting the ratio conditions 1:2 (pairs 1:2 vs. 3:6) and 1:3 (pairs 1:3 vs. 2:6) for both CP-knowers [1:2 ratio: $t(55) = -0.53$, $SD = 8.0$, $p = .60$; 1:3 ratio: $t(55) = -0.36$, $SD = 7.8$, $p = .72$] and 3–5 knowers [1:2 ratio: $t(27) = 0.2$, $SD = 9.73$, $p = .88$; 1:3 ratio: $t(27) = 1.7$, $SD = 9.2$, $p = .09$]. Consistently, the additional segregation of incongruent trials into small and large value trials also did not provide evidence for any absolute value effects for the P_{inc} found

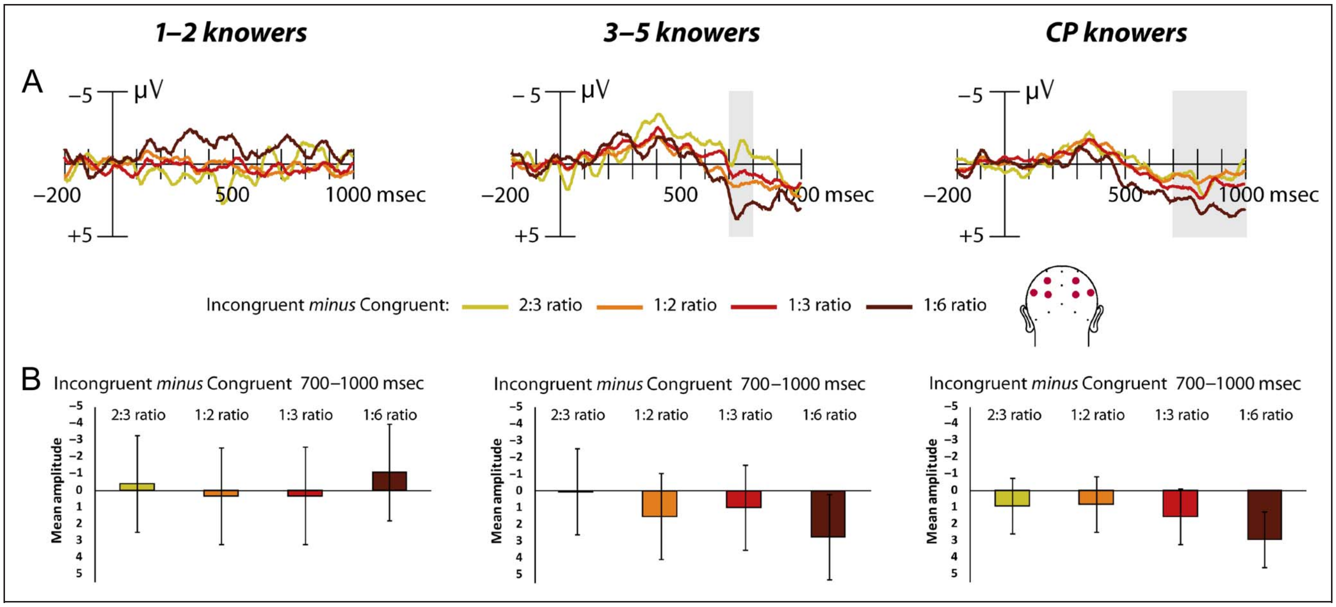

Figure 4. The P_{inc} (700–1000 msec) modulation by numerical ratio for 1–2 knowers (left) 3–5 knowers (middle), and CP-knowers (right): (A) Difference waves (incongruent minus congruent) averaged across the six highlighted channels for the 2:3, 1:2, 1:3, and 1:6 numerical ratio conditions. Traces are time-locked to the onset of the spoken number word. Each tick mark represents 100 msec. (B) Mean amplitude values (microvolts) of congruency from 700 to 1000 msec. Error bars denote .95 confidence intervals. The P_{inc} significantly increased with the increase in the numerical mismatch for both CP-knowers (700–1000 msec) and 3–5 knowers (700–800 msec).

Table 2. Correlations between Children's Age, Behavioral and ERP Measures

Pairs of Variables	*Spearman's Rho*	*Pearson's r*	*p*
N_{inc} - Knower level	−.11		.24
N_{inc} - DVAP		−.07	.43
N_{inc} - Age		−.06	.54
P_{inc} - Knower level	.21		.02
P_{inc} - DVAP		.06	.49
Pinc - Age		.08	.39
Knower level - DVAP	.51		.000001

N_{inc} and P_{inc} were evaluated as the mean amplitudes derived from the incongruent minus congruent difference waves. Knower level represents the highest number word (from one to six) a child knew. DVAP scores represent parental assessment of the child's vocabulary size.

in 3–5 knowers and CP-knowers. Although congruent trials significantly (or marginally) differed from incongruent small-value and large-value trials (ps ≤ .06), the two types of incongruent trials did not differ from one another (ps ≥ .15; Table 1). No other significant interactions with congruency were obtained in the P_{inc} analysis.

Finally, we tested whether 1–2 knowers showed N_{inc} and P_{inc} effects when the analysis was limited to the number words "one" and "two" and to visual arrays of one and two objects (i.e., "one" paired with a picture of 1 object and "two" paired with a picture of 2 objects vs. "one" paired with a picture of 2 objects and "two" paired with a picture of 1 object). As before, the results of this analysis did not reveal significant incongruency effects [N_{inc}: $t(30) = -0.54$, $SD = 6.5$, $p = .59$; P_{inc}: $t(30) = 0.85$, $SD = 6.8$, $p = .40$].

Correlational Analysis

Table 2 presents the Pearson's r and Spearman rank order correlations between the various ERP components and behavioral measures. As can be seen, individual mean amplitude values of the P_{inc} significantly correlated with knower level such that children who knew more number words showed larger P_{inc} amplitudes. Furthermore, knower level significantly correlated with DVAP scores such that children who knew more number words also had a larger vocabulary in general. Also of note is that children's age did not correlate with their N_{inc} or the P_{inc}. No other significant correlations emerged.

DISCUSSION

This study explored the neural bases of number word comprehension and its development in young children. We first asked whether children would show incongruency-related ERPs to spoken number words that were incongruent versus congruent with concurrently displayed arrays of objects. A second goal was to examine the brain activity associated with children's number word knower level and whether these data could speak to the debate over which nonverbal representations children first rely on when learning the meaning of number words. Children who showed the most limited number word knowledge (i.e., 1–2 knowers) showed no ERP incongruency effects. In contrast, two ERP components differentiated congruent from incongruent trials in children with intermediate knowledge level (3–5 knowers) and in children who comprehended the cardinal meaning of all the presented number words (CP-knowers): an early, central-parietal, negative polarity effect (N_{inc}) from 200 to 500 msec and a longer-latency, parietal, positive polarity effect (P_{inc}) from 700 to 1000 msec. Accordingly, children who knew more number words showed stronger incongruency-related ERPs, validating the use of the incongruency-related brain signature as a metric of children's semantic numerical knowledge (see also Henderson et al., 2011; McCleery et al., 2010; Friedrich & Friederici, 2004; Byrne et al., 1999).

The earliest effect, a negative polarity incongruency (N_{inc}), elicited from about 200 to 500 msec after the onset of the spoken number word had a central-parietal topography, similar to the N400 effect found in situations of semantically violated versus expected contexts (e.g., Henderson et al., 2011; McCleery et al., 2010; Friedrich & Friederici, 2004; Byrne et al., 1999; for a review, see Kutas & Federmeier, 2011). This central ERP negativity also resembled the N450/N_{inc} effects observed in the classical and the numerical versions of the Stroop task tested both in the visual (e.g., Szücs & Soltesz, 2007, 2012; Appelbaum et al., 2009; Larson et al., 2009; Szücs et al., 2009; West, 2003; Liotti et al., 2000) and auditory modalities (Donohue et al., 2012).

In our view, the N_{inc} effect observed here appears to be more related to conflict detection, possibly arising in part from the ACC, rather than to difficulties in carrying out semantic integration into context. Previous studies have linked the N450 to conflict effects arising from ACC, as modeled by source localization analyses (Szücs et al., 2009; West, 2003; Liotti et al., 2000). Consistent with this view are fMRI studies indicating that the anterior cingulate and other superior midline frontal regions play a role in stimulus conflict detection (e.g., Melcher & Gruber, 2009; Van Veen & Carter, 2005) as well as combined ERP/electromyography studies suggesting that the N450 is related to stimulus and not response conflict (e.g., Szücs & Soltesz, 2012; Szücs et al., 2009). Our interpretation is supported by the fact that the N_{inc} was not modulated by numerical ratio even in CP-knowers and was also not correlated with children's knower level. This interpretation is also in agreement with a previous numerical Stroop study in adults (Szücs et al., 2009) that found no sensitivity in the N450 amplitude to numerical distance. Yet, considering the differences in the N_{inc} distribution between the

3–5 knowers and the CP-knowers, it remains possible that the effect have somewhat different sets of neural sources. Further studies with source localization and/or fMRI measures are thus needed to determine the precise neural sources of this activity.

With regard to the latency of the N_{inc}, it seemed to onset here ~150 msec earlier than the N450 in classical visual Stroop tasks (e.g., Szücs et al., 2009; Liotti et al., 2000), but was similar to timing observed in an auditory version of the Stroop task (Donohue et al., 2012). Donohue et al. (2012) suggested that these latency differences might be partially related to faster sensory processing in the auditory modality. This possibility is consistent with an adult study by Szücs and Csepe (2005) that found early differences (<250 msec) in ERPs on fronto-central electrodes to auditory number words versus visually presented numbers.

Children who exhibited intermediate and high levels of number word knowledge also showed a sustained, positive-going deflection (P_{inc}) from 700 to 1000 post-stimulus, distributed over parietal sites. This effect was similar in topography to the SP incongruency effect reported to follow the N450 in adult studies of the Stroop task (e.g., Szücs & Soltesz, 2007, 2012; Appelbaum et al., 2009; Larson et al., 2009; West, 2003; Liotti et al., 2000), although its onset here was ~200 msec later. The functional role of the SP has not been clearly established. Some accounts suggest that it is related to late control processes relevant in tasks that require responses (Donohue et al., 2012; Larson et al., 2009); however, the current study involved children passively attending to multisensory stimuli, without any instruction to make an overt response. Another interpretation suggests that the effect might reflect additional semantic processing of word meaning given its tendency to be more prominent over left hemisphere sites and its proximity to Wernicke's area (Liotti et al., 2000). This explanation seems more probable with regard to the P_{inc} observed here, particularly because the effect was modulated by the degree of numerical incongruency; however, we found no statistical evidence for any lateralization of the P_{inc}. In our view and consistent with semantic-related interpretations of late posterior positive waves (e.g., Munte et al., 1998; Juottonen et al., 1996), the P_{inc} observed in this study seems likely to be associated with semantic processing of the multisensory numerical mismatch. First, the P_{inc} positively correlated with children's number word comprehension level. Second, children who showed mastery of the cardinal principle and children classified as 3–5 subset-knowers exhibited a P_{inc} that was modulated by the ratio of the numerical mismatch, suggesting that, unlike the N_{inc}, the P_{inc} is sensitive to the degree of semantic conflict. Although the obtained modulation of the P_{inc} by numerical ratio is the first to show ratio dependence in the neural processing of spoken number words in young children, it is consistent with many studies that have found ratio and distance ERP effects, usually elicited over posterior parietal scalp sites, for other numerical stimuli, in adults, children, and infants in diverse tasks (e.g., Soltesz et al., 2011; Szücs et al., 2007; Paulsen et al., 2010; Hyde & Spelke, 2009; Libertus et al., 2007; Szücs & Soltesz, 2007; Berger et al., 2006; Szücs & Csepe, 2005; Temple & Posner, 1998; Dehaene, 1996). Third, the fact that the P_{inc} was modulated by numerical ratio but not by the absolute numerical values suggests that children were processing multimodal numerical relationships consistent with engagement of the ANS.

Given that our study tracked the brain correlates associated with children's number word knower level, it provides the first neural data that can speak to the ongoing debate regarding whether approximate number representations or object file representations may serve as the basis for children's acquisition of number words (e.g., Carey, 2009; Sarnecka & Lee, 2009; Le Corre & Carey, 2007; Gelman & Gallistel, 2004; Pica et al., 2004; Gallistel & Gelman, 1990). We found that children who have already mastered their count list show robust modulation of the P_{inc} as a function of ratio with no sensitivity to the absolute numerical values. Furthermore, we also found such a ratio dependence in the P_{inc} from 700 to 800 msec in 3–5 knowers. Note that the numerical ratio effect found for these children, as well as both the N_{inc} and P_{inc} main effects, were all evidenced for the number words "one," "two," "three," and "six," that is, three number words that these children have already mastered and one number word that they still do not know according to their behavioral responses. The finding that 3–5 knowers show neural evidence of ratio-dependent processing of number words is particularly noteworthy because it suggests that number words are being mapped onto ANS representations before children have mastered the meaning of the number words within their count list. Accordingly, our results are consistent with the view that the ANS may serve as the foundation upon which number words are mapped, as it seems that ANS representations come into play before children acquire the cardinal principle (e.g., Wagner & Johnson, 2011; Gelman & Gallistel, 2004; Pica et al., 2004; Gallistel & Gelman, 1990). Our data do not, however, rule out the alternative hypothesis that children may first map number words onto object file representations and only later map number words onto the ANS (e.g., Sarnecka & Lee, 2009; Le Corre & Carey, 2007; Slaughter et al., 2006; Rousselle et al., 2004; Huntley-Fenner & Cannon, 2000). Future work would be necessary to directly test this hypothesis.

The more limited ratio dependence observed in 3–5 knowers (i.e., significant only from 700 to 800 msec and marginal from 800 to 900 msec) may be attributed to the fact that their number representations are much fuzzier than those of CP-knowers, such that larger ratios than those used in the current design would be needed to observe more robust ratio effects. Alternatively, the passive nature of our task design might have also contributed to the differences between knower level groups.

If children had been asked to actively detect multisensory mismatches, 3–5 knowers may have shown more robust ratio dependence. Future work is needed to test whether tasks that require children to provide overt responses and that include a larger set of number words would show more robust ratio-dependent ERPs for number words that are beyond children's current cardinal mastery. Such work may also provide better understanding with regard to the neural processing of number words in children who have very limited number word knowledge (e.g., one- and two-knowers). Furthermore, given the passive nature of our task design, we cannot rule out the possibility that the differences found between the ERPs of the various knower level groups might reflect in part the differences in children's counting skills, rather than differences in their number word knowledge. Finally, although we found no relation between children's verbal ability or their age with their number incongruency ERPs, it is still possible that other variables such as executive control functions may have contributed to the observed differences between groups and might partly explain the absence of the effects in the 1–2 knowers.

In summary, the patterns observed in the current study for children with intermediate and high levels of number word knowledge are consistent with the idea that two processing stages were evoked by the multisensory numerical mismatch, with differential functional characteristics. First, the earlier latency centroparietal N_{inc} (from 200 to 500 msec) marks detection of a semantic conflict. Second, the longer-latency, parietal distributed P_{inc} (700–1000 msec) marks a subsequent evaluation of the degree of the multimodal numerical incongruency, reflected by the modulation of the P_{inc} by numerical ratio. By the time children have mastered the cardinal principal the P_{inc} shows clear modulation of ratio, providing neural evidence consistent with various prior behavioral studies that show that by this point children have mapped symbolic representations onto approximate representations of number. A more novel finding reflected by this later processing component is that, even before children have fully mastered the cardinal principal, the grounding of symbols onto approximate representations appears to have already begun.

Acknowledgments

We thank the children and parents for their participation as well as David Paulsen, Melissa Libertus, Ken Roberts, and Celia Litovsky for helpful suggestions at various stages of this study; Vanessa Bermudez for assistance with data collection and analysis; and Tiffany Williamson, Krishan Sivaraj, Kristin Duffy, Stefanie Schwartz, Anna-Beth Keith, and Baninder Kaur for their assistance in data collection. This research was supported by grants from NIH (RO1-HD059108-06) and the McDonnell Foundation (Scholar Award) to E. M. B.

Reprint requests should be sent to Michal Pinhas, Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, or via e-mail: michalpinhas@gmail.com.

Notes

1. We also tested children's ability to count aloud up to 10, but performance in this task did not yield enough variance to be included in the analyses.
2. Note that only the relevant subset of congruent number words was used for the computation of the difference wave (incongruent minus congruent) for the ratio conditions 2:3 and 1:6. For example, when considering the incongruent pair 1:6 (i.e., trials with 1 object and the word "six" or 6 objects and the word "one"), only trials with congruent pairings of one and six (i.e., trials with 1 object and the word "six" or 6 objects and the word "one") were used for the subtraction.

REFERENCES

Appelbaum, L. G., Meyerhoff, K. L., & Woldorff, M. G. (2009). Priming and backward influences in the human brain: Processing interactions during the Stroop interference effect. *Cerebral Cortex, 19,* 2508–2521.

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

Brannon, E. M., Abbott, S., & Lutz, D. (2004). Number bias for the discrimination of large visual sets in infancy. *Cognition, 93,* B59–B68.

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1-9 by monkeys. *Science, 282,* 746–749.

Byrne, J. M., Connolly, J. F., MacLean, S. E., Dooley, J. M., Gordon, K. E., & Beattie, T. L. (1999). Brain activity and language assessment using event-related potentials: Development of a clinical protocol. *Developmental Medicine and Child Neurology, 41,* 740–747.

Carey, S. (2009). *The origin of concepts*. Cambridge, MA: MIT Press.

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract?. *Behavioral and Brain Sciences, 32,* 313–328.

Condry, K. F., & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General, 137,* 22–38.

Cordes, S., & Brannon, E. M. (2009). Crossing the divide: Infants discriminate small from large numerosities. *Developmental Psychology, 45,* 1583–1594.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8,* 698–707.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience, 8,* 47–68.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences, 21,* 355–361.

Donohue, S. E., Liotti, M., Perez, R., III, & Woldorff, M. G. (2012). Is conflict monitoring supramodal? Spatiotemporal dynamics of cognitive control processes in an auditory Stroop task. *Cognitive, Affective and Behavioral Neuroscience, 12,* 1–15.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio, TX: Pearson.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science, 8,* 307–314.

Friedman, D., Sutton, S., Putnam, L., Brown, C., Jr., & Erlenmeyer-Kimling, L. (1988). ERP components in picture matching in children and adults. *Psychophysiology, 25,* 570–590.

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

Gallistel, C. R., & Gelman, R. (1990). The what and how of counting. *Cognition, 34,* 197–199.

Gelman, R., & Gallistel, C. R. (2004). Language and the origin of numerical concepts. *Science, 306,* 441–443.

Henderson, L. M., Baseler, H. A., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language, 117,* 88–99.

Hsu, Y.-F., & Szücs, D. (2010). The time course of symbolic number adaptation: Oscillatory EEG activity and event-related potential analysis. *Neuroimage, 59,* 3103–3109.

Huntley-Fenner, G., & Cannon, E. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science, 11,* 147–152.

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: An electrophysiological investigation of small and large number representations. *Journal of Cognitive Neuroscience, 21,* 1039–1053.

Juottonen, K., Revonsuo, A., & Lang, H. (1996). Dissimilar age influences on two ERP waveforms (LPC and N400) reflecting semantic context effect. *Cognitive Brain Research, 4,* 99–107.

Kok, A., & Rooijakkers, J. A. (1985). Comparison of event-related potentials of young children and adults in a visual recognition and word reading task. *Psychophysiology, 22,* 11–23.

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.

Larson, M. J., Kaufman, D. A. S., & Perlstein, W. M. (2009). Neural time course of conflict adaptation effects on the Stroop task. *Neuropsychologia, 47,* 663–670.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105,* 395–343.

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52,* 130–169.

Libertus, M. E., & Brannon, E. M. (2009). Behavioral and neural basis of number sense in infancy. *Current Directions in Psychological Science, 18,* 346–351.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the Approximate Number System correlates with school math ability. *Developmental Science, 14,* 1292–1300.

Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2013). A Developmental Vocabulary Assessment for Parents (DVAP): Validating parental report of vocabulary size in 2-7-year-old children. *Journal of Cognition and Development*. doi:10.1080/15248372.2013.835312.

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions, 3,* 1–15.

Licht, R., Kok, A., Bakker, D. J., & Bouma, A. (1988). The development of lateral event-related potentials (ERPs) related to word naming: A four year longitudinal study. *Neuropsychologia, 26,* 327–340.

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of Stroop color-word interference effect. *Neuropsychologia, 38,* 701–711.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large number discrimination in human infants. *Psychological Science, 15,* 396–401.

McCleery, J., Ceponiene, R., Burner, K. M., Townsend, J., Kinnear, M., & Schreibman, L. (2010). Neural correlates of verbal and nonverbal semantic integration in children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 51,* 277–286.

Melcher, T., & Gruber, O. (2009). Decomposing interference during Stroop performance into different conflict factors: An event-related fMRI study. *Cortex, 45,* 189–200.

Munte, T. F., Heinze, H. J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36,* 217–226.

Paulsen, D. J., Woldorff, M. G., & Brannon, E. M. (2010). Individual differences in nonverbal number discrimination correlate with event-related potentials and measures of probabilistic reasoning. *Neuropsychologia, 48,* 3687–3695.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306,* 499–503.

Rebai, M., Bernard, C., & Lannou, J. (1997). The Stroop's test evokes a negative brain potential, the N400. *The International Journal of Neuroscience, 91,* 85–94.

Rousselle, L., Palmers, E., & Noel, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology, 87,* 57–84.

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition, 108,* 662–674.

Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology, 103,* 325–337.

Sitnikova, T., Kuperberg, G., & Holcomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology, 40,* 160–164.

Slaughter, V., Kamppi, D., & Paynter, J. (2006). Toddler subtraction with large sets: Further evidence for an analog-magnitude representation of number. *Developmental Science, 9,* 33–39.

Soltesz, F., White, S., & Szücs, D. (2011). Event-related brain potentials dissociate the developmental time-course of automatic numerical magnitude analysis and cognitive control functions during the first three years of primary school. *Developmental Neuropsychology, 36,* 682–701.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 18,* 643–662.

Szücs, D., & Csepe, V. (2005). The parietal distance effect appears in both the congenitally blind and matched sighted controls in an acoustic number comparison task. *Neuroscience Letters, 384,* 11–16.

Szücs, D., & Soltesz, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia, 45,* 3190–3202.

Szücs, D., & Soltesz, F. (2012). Functional definition of the N450 event-related brain potential marker of conflict processing: A numerical Stroop study. *BioMed Central Neuroscience, 13,* 35.

Szücs, D., Soltesz, F., Jarmi, E., & Csepe, V. (2007). The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: An electro-encephalography study. *Behavioral and Brain Functions, 23,* 1–20.

Szücs, D., Soltesz, F., & White, S. (2009). Motor conflict in Stroop tasks: Direct evidence from single-trial electro-myography and electro-encephalography. *Neuroimage, 47,* 1960–1973.

Temple, E., & Posner, M. I. (1998). Brian mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences, U.S.A., 95,* 7836–7841.

Van Veen, V., & Carter, S. C. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional fMRI study. *Neuroimage, 27,* 497–504.

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition, 119,* 10–22.

West, R. (2003). Neural correlates of cognitive control and conflict detection in the Stroop and digit location tasks. *Neuropsychologia, 41,* 1122–1135.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155–193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24,* 220–251.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74,* B1–B11.