



Published in final edited form as:

Dev Sci. 2019 January ; 22(1): e12746. doi:10.1111/desc.12746.

The organization of words and environmental sounds in the 2nd year: Behavioral and Electrophysiological evidence

Kristi Hendrickson¹, Tracy Love^{2,4}, Matthew Walenski⁵, and Margaret Friend³

¹Department of Communication Sciences & Disorders, University of Iowa, USA

²Center for Research in Language, University of California, San Diego, USA

³Department of Psychology, San Diego State University, USA

⁴School of Speech, Language, and Hearing Sciences, San Diego State University, USA

⁵School of Communication, Northwestern University, USA

Abstract

The majority of research examining early auditory-semantic processing and organization is based on studies of meaningful relations between words and referents. However, a thorough investigation into the fundamental relation between acoustic signals and meaning requires an understanding of how meaning is associated with both lexical and non-lexical sounds. Indeed, it is unknown how meaningful auditory information that is not lexical (e.g., environmental sounds) is processed and organized in the young brain. To capture the structure of semantic organization for words and environmental sounds, we record event-related potentials (ERPs) as 20-month-olds view images of common nouns (e.g., dog) while hearing words or environmental sounds that match the picture (e.g., “dog” or barking), that are within-category violations (e.g., “cat” or meowing), or that are between-category violations (e.g., “pen” or scribbling). Results show both words and environmental sounds exhibit larger negative amplitudes to between-category violations relative to matches. Unlike words, which show a greater negative response early and consistently to within-category violations, such an effect for environmental sounds occurs late in semantic processing. Thus, as in adults, the young brain represents semantic relations between words and between environmental sounds, though it more readily differentiates semantically similar words compared to environmental sounds.

1. Introduction

Auditory-semantic knowledge requires an appreciation of the relation between sounds and concepts, and an understanding of how concepts relate to one another. Indeed our ability to interpret the world depends fundamentally on how the brain organizes meaningful auditory information. In adults, lexical-semantic information exhibits a fine-grained organizational structure based on featural similarity – the perceived likeness between concepts – which aids

Corresponding Author, Kristi Hendrickson, Department of Psychology, University of Iowa, Kristi-hendrickson@uiowa.edu, (925) 323-3763; Tracy Love, School of Speech, Language, and Hearing Sciences, San Diego State University, tlove@mail.sdsu.edu; Matthew Walenski, School of Communication, Northwestern University, matthew.walenski@northwestern.edu; Margaret Friend, Department of Psychology, San Diego State University, mfriender@mail.sdsu.edu.

in categorization (Kay, 1971; Murphy, Hampton, & Milovanovic, 2012; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Sajin, & Connine, 2014; Hendrickson, Walenski, Friend, & Love, 2015). There is recent evidence that the organization of the early lexical-semantic system may be mediated by featural information similar to the organization observed in adults (Ariejas-Trejo & Plunkett, 2009; 2013; Borovsky, Ellis, Evans, & Elman, 2015; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willits, Wojcik, Seidenberg, & Saffran, 2013).

To date, models of early auditory-semantic processing are primarily based on studies of meaningful relations between words and referents. However, auditory-semantic information can be divided into two categories: lexical (i.e., words) and non-lexical (e.g., environmental sounds such as the sound of a dog barking). Unlike words, which bear an arbitrary relation to real-world referents, and thus vary across languages, environmental sounds have an inherent correspondence to a visual referent (Van Petten & Rheinfelder, 1995). Therefore, a thorough investigation into the fundamental relation between acoustic signals and meaning requires an understanding of how meaning is associated with both lexical and non-lexical sounds.

Such an investigation can further our understanding of the relation between language and cognition by examining whether an interconnected auditory-semantic network can be instantiated independent of language early in development. What's more, it has recently been suggested that the consistency with which environmental sounds are associated with their object referents may bootstrap the learning of more arbitrary word-object relations (Cummings et al., 2009). However, this claim is based on the assumption that the mechanisms of semantic integration that subservise the processing of words and environmental sounds are similar in the developing brain. Event-related potentials (ERPs) are a brain-based method that can identify well-defined stages of meaningful auditory processing and have been used to explore neural correlates of auditory-semantic integration and organization. The overarching objective of this study is to use behavioral measures and ERPs to compare how words and environmental sounds are organized at the semantic level early in language development.

1.1 Word vs. environmental sound processing in adults

The N400, a negative wave peaking approximately 400 ms post-stimulus onset, is an ERP component closely tied to semantic processing (Kutas & Federmeir, 2011; Kutas & Hillyard, 1980). All semantic stimuli (auditory, pictorial, orthographic) elicit an N400, whose amplitude is larger when the stimulus violates an expectancy set by a preceding semantic context. The N400 incongruity effect denotes the relative increase in N400 amplitude to a semantically unrelated stimulus.

N400 incongruity effects have been found for words or pictures primed by related and unrelated environmental sounds (Schön, Ystad, Kronland-Martinnet, & Besson, 2010; Daltrozzo & Schön, 2009; Frey, Aramaki, & Besson, 2014; Van Petten & Rheinfelder, 1995) and for environmental sounds primed by related and unrelated words, pictures, or other environmental sounds (Aramaki, Marie, Kronland-Martinnet, Ystad, & Besson, 2010; Cummings et al.; 2006; 2008; 2010; Daltrozzo & Schön, 2009; Orgs, Lange, Dombrowski,

& Heil, 2008; Orgs et al., 2006; Plante, Van Petten, & Senkfor, 2000; Schirmer, Soh, Penney, & Wyse, 2011; Schön et al., 2010; Van Petten & Rheinfelder, 1995).

Not only is the N400 sensitive to semantic congruency (Kiefer, 2001; Kutas & Hillyard, 1980; Nigam, Hoffman, & Simons, 1992), but it is also incrementally sensitive to differences in the featural similarity of concepts to which words refer. Specifically, both within- and between-category violations exhibit significant N400 effects; however, between-category violations (e.g., ‘jeans’ instead of ‘boots’) exhibit greater N400 amplitudes than within-category violations (e.g., ‘sandals’ instead of ‘boots’) (Federmeier & Kutas 1999b; 2002; Federmeier, McLennan, Ochoa, & Kutas, 2002; Ibanez, Lopez, & Cornejo, 2006). A previous study of adults reported that words exhibited graded N400 amplitudes starting early (300 ms) and this graded structure continued throughout all time windows of interest. Conversely, for environmental sounds, only between-category violations (not within-category violations) exhibited N400 effects in the early time window, although a graded pattern similar to that of words was exhibited at a later latency time window (400 – 500 ms) (Hendrickson et al., 2015). These results indicate that for adults, the organization of words and environmental sounds in memory is differentially influenced by featural similarity, with a consistently fine-grained graded structure for words but not sounds.

1.2 Lexical-semantic organization in young children

Recent work examining when and how infants develop a system of words that are semantically related comes from studies that use infant adaptations of adult lexical priming paradigms (Ariejas-Trejo & Plunkett, 2009; 2013; Delle Luche, Durrant, Floccia, & Plunkett, 2014; Hendrickson & Sundara, 2016; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willits, et al., 2013). Two-year-olds demonstrate a lexical priming effect such that related word primes yield longer looking times to a visual referent relative to unrelated primes (Arias-Trejo & Plunkett, 2009; 2013). This suggests that by their 2nd year, children organize their lexical network based on the associative and featural similarities among semantic referents.

The neural architecture that underlies the N400 response develops ontogenetically early, as N400 effects have been observed at 6-months to word-picture violation and at 9-months to unanticipated action sequences (Federmeier & Kutas, 2011; Friedrich, Wilhelm, Born, & Friederici, 2015; Reid, Hoehl, Grigutsch, Groendahl, Parise, & Striano, 2009). In the domain of language, N400 effects in response to picture-word violations – such as those employed in the current study – have been seen as young as 12 -14 months (Friedrich & Friederici, 2008). Research within the last decade has consistently shown that children in their second year exhibit an N400-like incongruity response to a variety of lexical-semantic violations. (Friedrich & Friederici, 2004, 2005; 2008; 2015; von Koss Torkildsen et al., 2006; 2009; Mills, Conboy, & Paton, 2005; Rama, Sirri, & Serres 2013). What’s more, 20-month-olds display an N400-like incongruity effect that changes as a function of category membership; the incongruity response is earlier and larger for between-category violations (e.g., dog and chair) than within-category violations (e.g., dog and cat) (Koss, Torkildsen, et al., 2006).

1.3 Lexical vs. non-lexical auditory processing in young children

Early in development (< 10 months), infants have a preference for listening to verbal compared to nonverbal sounds (e.g., tones, animal calls; Krentz & Corina, 2008; Vouloumanos & Werker, 2004; 2007). Research on the *semantic* processing of verbal vs. nonverbal sounds suggests toddlers younger than 24 months use both types of auditory information to differentiate concepts based on their degree of featural similarity (Campbell & Namy, 2003; Hollich et al., 2000; Namy, 2001; Namy, Acredolo, & Goodwyn, 2000; Namy & Waxman, 1998, 2002; Roberts, 1995; Sheehan et al., 2007), whereas toddlers 24 months and older use verbal, but *not* nonverbal information to differentiate concepts based on their degree of featural similarity (Namy & Waxman, 1998; Sheehan, Namy, Mills, 2007). Thus, it has been suggested that at 24 months, children undergo a form of linguistic specialization in which they begin to understand the role that language – as opposed to other types of information – plays in organizing objects according to subtle differences in featural similarity (Namy, Campbell, & Tomasello, 2004). However, these studies used sounds (e.g., tones) that lack a natural semantic association with the paired visual referent, and thus are unlikely to tap into auditory-semantic systems. Therefore, these studies do not directly assess toddler’s understanding of lexical versus inherently *meaningful* non-lexical sounds (e.g., environmental sounds) as a means of conveying semantic information.

Only one study has examined the semantic processing of environmental sounds younger than age seven. Cummings, Saygin, Bates, and Dick (2009) tested 15-, 20-, and 25-month-olds’ using a looking-while-listening paradigm. Participants heard environmental sounds or spoken words when viewing pairs of images and eye movements to match versus non-match pictures were captured to determine the accuracy of object identification. Object recognition for environmental sounds and words was found to be strikingly similar across ages.

Although Cummings et al. (2009) found recognition of sound-object associations for environmental sounds and words to be quite similar throughout the 2nd year of life, such results cannot shed light on whether there are differences in the underlying processes driving infants’ overt responses to words vs. environmental sounds. Nor can the extant literature illuminate whether there are differences in how words and environmental sounds are semantically organized in the brain early in development.

1.4 Current Study

In the current study, we investigate the structure of semantic organization for words and environmental sounds in the 2nd year of life by using ERPs. We chose to investigate 20-month-olds for two reasons. First, it has been suggested that at 24 months the processing of words undergoes specialization, such that words become the predominant form of labeling semantically different but related items. Specifically, behavioral work suggests 24-month-olds begin to understand the role that language – as opposed to other types of information – plays in organizing objects according to subtle differences in featural similarity (Namy, Campbell, & Tomasello, 2004; Sheehan, Namy, & Mills, 2007). Second, 20 months is the youngest that a semantically graded lexicon has been observed in children (Torkildson et al., 2006), such that violations with greater semantic feature overlap (e.g., “cat” when “dog” is expected) exhibit attenuated N400 effects compared to violations with less semantic feature

overlap (e.g., “chair” when “dog” is expected). Therefore 20-month-olds were chosen to ensure replication of previous findings for words (Torkildson et al., 2006), and to compare the organization of words and environmental sounds before and after (see Hendrickson et al., 2015 for adult findings) this putative shift in linguistic specialization is said to occur.

In order to ensure that any observed differences in organizational structure between words and environmental sounds were not due to *a priori* exposure, familiarity or comprehension, a language and environmental sounds familiarization and assessment session was conducted within one and a half weeks prior to participation in the electrophysiological session. First, 20-month-olds participate in a behavioral familiarization task, to ensure that each child is familiar with both the words and environmental sounds associated with each concept tested during an ERP task. Subsequently, they participate in a picture-pointing task to assess word and environmental sound comprehension and speed of processing. Finally, we assess the organizational structure of words and environmental sounds by recording ERPs as participants view images of common nouns (e.g., dog) while hearing words or environmental sounds that match the picture (e.g., “dog” or barking), those that are a within-category violations (e.g., “cat” or meowing), and those that are a between-category violations (e.g., “pen” or scribbling).

This study has two aims. First, we examine whether children in their 2nd year show N400 incongruity effects for environmental sounds preceded by pictures that constitute within- or between-category semantic violations. Toddlers’ ERP response to environmental sounds has not yet been examined. Therefore, the onset of the N400 – an ERP component linked to semantic integration – to this type of stimulus is unknown. We seek to establish whether N400 incongruity effects are obtained for environmental sounds given egregious semantic violations (between-category violation) since this is the contrast most likely to yield such an effect. Behavioral evidence suggests that children this age show similar performance in their recognition of familiar words and environmental sounds when presented with between category pictures (15-25-months; Cummings et al., 2009). Thus, we expect environmental sounds, like words, to exhibit significant N400 effects to between-category violations at 20 months. Further, we expect children will show similar performance in their ability to comprehend referents for words and environmental sounds during the picture-pointing task.

The second aim of the study is to compare semantic organization of words and environmental sounds at 20 months. This aim primarily concerns how the N400 amplitude of within-category violations compares relative to the matches and between-category violations for each sound type. If at 24 months linguistic information undergoes specialization such that words become the predominant form of organizing semantically different but related items, we expect to observe different patterns of brain activity for each sound type at 20 months compared to adults (Hendrickson et al., 2015). Specifically, we expect that for 20-month-olds condition (between or within) will influence the ERP response for words and environmental sounds similarly: Both violations will exhibit significant N400 effects, however between-category violations will exhibit greater N400 amplitudes than within-category violations. However, like adults, environmental sounds could be organized more coarsely than words early in language development. Therefore, an alternative prediction is that the processing of words and environmental sounds will be differentially

influenced by the degree of semantic violation, with an earlier and more consistent fine-grained structure apparent for words than sounds. From this view, we expect the ERP response to within-category violations to differ from matches and we expect this difference to be more robust (i.e., start earlier and last longer and/or evince larger effect sizes) for words compared to environmental sounds.

2. Materials and Methods

2.1 Participants

Children were obtained through a database of parent volunteers recruited through birth records, internet resources, and community events in a large metropolitan area. Only those children with at least 80% language exposure to English were included in the study (Bosch & Sebastian-Galles, 2001). All infants were full-term and had no diagnosed impairments in hearing or vision. Overall, 26 children participated in this within-subjects study. All 26 participants completed the behavioral familiarization task. Attention during the behavioral familiarization task was decided by examiner's observations and post-hoc determination of percentage of trials completed. If the participant failed to look at the screen on three consecutive trials given attempts by the experimenter to re-engage testing was terminated. The final sample was attentive (looking at the screen) for 90% of trials.

Of 26 children who participated in the behavioral familiarization task, 18 children (8 F; 10 M) with a mean age of 20.5 months (range = 19.5 – 21.5 months) were included in the final analysis related to the picture-pointing task; eight participants were excluded due to failure to complete the picture-pointing task. The criteria for ending testing for the picture-pointing task were taken from the protocol of the well-documented Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, Friend, Schmitt, & Simpson, 2012). Specifically, the task was ended if the participant failed to touch on two consecutive trials with two attempts by the experimenter to re-engage without success. If the attempts to re-engage were unsuccessful and the child was fussy, the task was terminated and the responses up to that point were taken as the final score. However, if the child did not touch for two or more consecutive trials but was *not* fussy, testing continued.

The ERP study was applied to 25 of the original 26 children (one child did not return for the EEG session of the study). Of these 25, six children were excluded from the ERP study because of refusal to wear the cap ($n = 2$), and failure to obtain at least 10 artifact-free trials per condition for either words or environmental sounds ($n = 4$). The final sample for the ERP study included 19 monolingual children (9 F; 10 M; mean age = 20.6 months, range = 19.5 – 21.5 months). The final within-subjects sample (including both the behavioral and ERP tasks) was 16 children (7 F; 9 M).

2.2 Stimuli

Stimuli for the behavioral and ERP tasks were colorful line drawings, auditory words, and environmental sounds of 30 highly familiar concepts (all of which were nouns). Concepts fall into one of three categories: animals (dog, cat, owl, sheep, horse, cow, bird, frog, bee, elephant, duck, bear, chicken, monkey, pig), vehicles (fire truck, car, train, bicycle,

motorcycle, airplane), and household objects (hammer, door, telephone, pen, clock, toothbrush, keys, zipper, broom). A female native English speaker produced the word stimuli (mean duration = 873 ms, SD = 200 ms), which were recorded in a single session in a sound-attenuating booth (sampling at 44.1 Hz, in 16-bit stereo). The average pitch of the word stimuli was 264.50 Hz (SD = 41.1 Hz). Environmental sound stimuli were obtained from several online sources (www.soundbible.com, www.soundboard.com, and www.findsounds.com) and from a freely downloadable database of normed environmental sounds (Hocking, Dzafic, Kazovsky, & Copland, 2013). Environmental sounds were standardized for sound quality (44.1 kHz, 16bit, stereo) and had a mean duration of 878 ms (SD = 251), and a mean pitch of 221.99 (SD = 119.20). The duration and pitch of the word and environmental sound stimuli did not significantly differ (duration: $t(30) = .033$, $p = .97$; pitch: $t(30) = 1.84$, $p = .07$). Visual stimuli were colorful drawings taken from Snodgrass and Vanderwart (1980).

Although comprehension norms are not available for children at 20 months of age, the words for the concepts used in this study are comprehended by an average of 54% of 16-month-olds, and produced by 62% of 24-month-olds (Fenson et al., 1994a). Therefore, these concepts should be highly familiar to children of 20 months. It is noted, however, that comprehension norms are not available for environmental sounds in that study. In order to ensure that the concepts were associated with easily identifiable environmental sounds, a Likert scale pretest was conducted. Ten native English-speaking college undergraduates were presented with 51 images of prototypical members of highly familiar concepts (e.g., dog) paired with an associated environmental sound (e.g., *barking*). Participants were asked to rate, on a 1-5 scale (1= not related and 5= very/highly related), how well the picture and sound went together. Each image was presented twice, though in a randomized order, each time with a different exemplar of an associated environmental sound. Therefore, 102 presentations of image/sound pairs were presented one at a time. Only those sounds that received a mean rating of 3.5 or higher were included as stimuli. If both sounds for the same image were above 3.5, we chose the sound with the higher score; if both sounds obtained the same score, we chose the sound we thought was more stereotypical. This procedure resulted in a total of 30 items that met our criteria for inclusion. The same 30 items were used to make match, within-category, and between-category conditions, and therefore conditions were very well controlled for word frequency, imageability, concreteness, phonology, and other properties of the stimuli.

2.3 Procedure

2.3.1 Session 1. Language and Environmental Sound Familiarization and Assessment

—Within one and a half weeks prior to participation in the electrophysiological session, each subject participated in a language and environmental sound familiarization and assessment. First, the familiarization phase was used to ensure that each child was familiar with both the words and environmental sounds associated with each of the 30 concepts tested during the ERP task. Second, the assessment phase directly gauged participant's understanding of the words and environmental sounds using a two-alternative force-choice procedure. Third, we obtained parental ratings of participant's *a priori* familiarity with the words and environmental sounds associated with each concept. Together

this assessment allowed us to equate exposure to the picture, word, and environmental sound stimuli, and determine whether there exist differences in participant's comprehension of the word- and environmental sound-concept associations.

2.3.1.1 Parental Rating Scale: Here, parents are asked to rate their child's familiarity with the list of words and environmental sounds from which the stimuli used in the ERP and behavioral tasks are drawn (30 concepts in all). Similar to the rating scale used by Sheehan, Namy, & Mills (2007) for words and gestures, parents rate each item from 1 (certain their child is not familiar with it) to 7 (certain their child is familiar with it).

2.3.1.2 Behavioral Familiarization Task.: To help control for exposure and familiarity effects with the specific stimuli used in the ERP and behavioral tasks, we familiarize participants with the word-picture and environmental sound-picture combinations an equal number of times (total duration of task = 6 mins.). During this familiarization phase, participants are presented with each concept (30 in all) on a computer monitor, 6 times each, 3 with the corresponding environmental sound, and 3 with the corresponding word, in randomized order to equate *a priori* levels of exposure.

2.3.1.3 Picture-pointing Task: This task largely followed the protocol for the Computerized Comprehension Task (CCT; Friend & Keplinger, 2003, Friend, Schmitt, & Simpson, 2012). The CCT is two-alternative forced-choice touch screen task that measures early decontextualized word knowledge. Previous studies have reported that the CCT has strong internal consistency (Form A: $a = .836$; Form B: $a = .839$), converges with parent report (partial r controlling for age = $.361$, $p < .01$), and predicts subsequent language production (Friend et al., 2012). In addition, responses on the CCT are nonrandom (Friend & Keplinger, 2008) and this finding replicates across languages (Friend & Zesiger, 2011) as well as for monolinguals and bilinguals (Poulin-Dubois, Bialystok, Blaye, Polonia & Yott, 2013). In the present study, we use this as a manipulation check to insure that children's comprehension of words and environmental sounds was matched prior to ERP testing.

For this procedure, infants are prompted to touch images on a touchscreen monitor by an experimenter seated to their right. Target touches (e.g. touching the image of the dog) elicit congruous auditory feedback over audio speakers (e.g. the word "dog", or the sound of a dog barking). Participants see each target picture (30 images used in the EEG task) with a yoked, between-category distractor image twice, once to test word comprehension (word block) and once to test environmental sound comprehension (environmental sound block). Note, the same-yoked pairs were presented in the word and environmental sound blocks (i.e., target *dog* was presented with distractor image *ball*, in both the word and environmental sound blocks). The order of blocks was counterbalanced such that half the participants received the word block first. Each block contained two training trials to ensure that participants understood the nature of the task. If the child failed to touch the screen after repeated prompts, the experimenter touched the target image for them. If a participant failed to touch during training, the two training trials were repeated once. Only participants who executed at least one correct touch during the training phase proceeded to the testing phase. For a given trial, first an inter-trial gray screen was presented. Once the participant directed their gaze toward the grey screen, the experimenter delivered a sentence prompt in infant-directed

speech (*words*: Where is the ___?; *environmental sounds*: Which one goes ___?). Directly after the sentence prompt, the experimenter clicked the mouse to present the two images simultaneously on the right and left side of the screen, and 250 ms after the images appeared on the screen, the *computer* elicited the target word or environmental sound. The side on which the target image appeared was presented in pseudo-random order across trials such that the target was presented with equal frequency on both sides of the screen (Hirsh-Pasek & Golinkoff, 1996). Once the image pair was presented, each trial lasted until the infant touched the screen or until 7 seconds had elapsed at which point the image pair disappeared and the inter-trial grey screen appeared.

The criterion for ending testing was a failure to touch on two consecutive trials with two attempts by the experimenter to re-engage without success. If the attempts to re-engage were unsuccessful and the child was fussy, the task was terminated and the responses up to that point were taken as the final score. However, if the child did not touch for two or more consecutive trials but was not fussy, testing continued.

There were two measures obtained for this task: accuracy and reaction time. Accuracy was measured as the number of target touches executed during the task. Reaction time was calculated for target touch trials, and was measured at the moment the participant made contact with the touch screen upon hearing the target word or sound –i.e., the time from word or sound onset to touch response.

2.3.2 Session 2. Event-Related Potential Study—For the ERP study, sound class (word, environmental) was presented in a blocked fashion, resulting in two back-to-back runs with three conditions per run (match, within-category violation, between-category violation). Each of the runs was composed of a presentation list with 90 trials (30 trials for each condition). The presentation list was constructed so that a particular picture was not repeated on consecutive trials, and a particular sound was not repeated within 6 trials. Further, presentation of conditions was pseudo-randomized across the presentation list such that a given condition (match, within-category, between-category) did not appear for more than three consecutive trials.

2.3.2.1 ERP Testing Procedure: Participants were seated on their caregiver’s lap at a distance of roughly 140 cm from a LCD computer monitor in a dimly lit, electrically shielded and sound-attenuated room. Each subject participated in two back-to-back runs, one for each sound type, each lasting approximately 8 minutes. The only difference between the runs was the type of sound presented (word or environmental). The order of the runs was counterbalanced such that half the participants received the word run first. The same 30 concepts were used to make matches, within-category violations, and between-category violations conditions, thus each auditory stimulus appeared three times, resulting in a balanced design. As shown in Figure 1 (below), for each trial, participants were presented with a colorful drawing of a familiar concept. The pictures were centered on screen and relatively small, so that they could be identified by central fixation (subtending a visual angle of 4.95 degrees on average). After 1750 ms participants heard a sound from one of three conditions (match, within-category violation, between-category violation). The picture disappeared at the offset of the sound (460 - 1235 ms). An inter-trial interval grey screen

was then presented, its timing varied randomly between 500 and 1500 milliseconds. To maintain children's interest, an attention getter was programmed to appear on the screen every 10 trials and when the participant looked away from the screen for more than 2 seconds. Participants were video recorded during the EEG testing session to reject trials in which participants were not looking at the screen.

2.3.2.2 EEG Recording: EEG data was collected using a 21-electrode cap (Electro cap Inc.) according to the International 10-20 system. Tin electrodes were placed at the following locations (FP1, FP2, F7, F3, FZ, F4, F8, C3, CZ, C4, M1, M2, P3, PZ, P4, T3, T4, T5, T6, O1, O2) (see Figure 2). All channels were referenced to the left mastoid during data acquisition; data was re-referenced offline to the average of the left- and right-mastoid tracings. EEG was recorded at a sampling rate of 500 Hz, amplified with a Neuroscan Nuamps amplifier and low-pass filtered at 100hz. EEG gain was set to 20,000 and EOG gain set to 5,000. Electrode impedances were mostly below 5 K Ω , but at least below 20 K Ω .

2.3.2.3 EEG Analysis: EEG was time locked to the auditory stimulus onset (spoken word or environmental sound) and epochs of 1200 ms from auditory onset were averaged with a 200 ms pre-stimulus baseline. A zero-phase digital band-pass filter ranging from 0.2 to 30 Hz was applied to the EEG data. Before averaging, trials in which the child was not looking at the screen, and trials containing eye movements, blinks, excessive muscle activity, or amplifier blocking were rejected by off-line visual inspection of the EEG data and video recording. The average rejection rate was comparable between words (39.5%) and environmental sounds (36.9%). Participants were included in the final data set if they had 10 artifact-free trials per condition. Data for one subject in the Word run, and two subjects in the Environmental Sound run were removed due to insufficient data per condition (< 10 artifact-free trials). To analyze potential differences in distributional effects across conditions while minimizing the number of total comparisons, we identified three regions of interest (ROI): Frontal (Fp1, Fp2, F7, F3, Fz, F4, F8), Central-Parietal (C3, Cz, C4, P3, Pz, P4), and Temporal-Occipital (T3, T5, O1, O1, T6, T4). Pooling all electrodes into these regions allowed us to enhance the signal to noise ratio (i.e., increase statistical power), and analyze effects of anteriority, while preserving the electrode sites where N400 effects have been shown to be maximal for similarly aged participants (central-parietal sites)

Prior work indicates that N400 incongruity effects (i.e. unrelated items are more negative than related items), start earlier, and last longer in the auditory as opposed to the visual modality (Holcomb & Neville, 1990). Based on this prior work, and visual inspection of the grand average waveforms, four time windows of interest were chosen: 200 – 400 ms, 400 – 600 ms, 600 – 800 ms, and 800 – 1000 ms. For each sound type (word and environmental), mean amplitude voltage was computed separately for each condition (match, within-category, and between-category) and electrode site within the four time windows of interest. For each sound type we analyzed these mean amplitude voltages using restricted maximum likelihood in a mixed-effects regression model with a random effect of subject on the intercept, fit with an unstructured covariance matrix. The model also included Condition (match, within-category, between-category), ROI (Frontal, Central-Parietal, Temporal-Occipital) and their interactions. Condition was contrasted within each region, and we report

the regression coefficients (and standard errors), t -values, p -values, 95% confidence interval, and effect size for these contrasts (see Tables 3 & 4). For effect size, we report Hedge's g_{av} (Lakens, 2013). The use of regression models offers several advantages over traditional ANOVA models, including robustness to unbalanced designs and a flexible ability to model different covariance structures, avoiding the need to correct for sphericity violations (see Newman et al., 2012 and references therein).

3. Results

3.1 Session 1: Behavioral and Parental Rating Results

To be included in analyses related to the picture-pointing task and the ERP task participants were required to complete the behavioral familiarization task. Of the 26 children who participated in the study, all were quiet and alert and maintained attention toward the screen for the duration of the behavioral familiarization task (6 mins.). Parental ratings of word and environmental sound familiarity were collected on all but one participant ($N = 25$). There was a significant difference in parent reported familiarity with the words and environmental sounds ($t(24) = 5.67$, $p < .0001$), such that parents reported that their children would be more familiar with the word stimuli compared to the environmental sound stimuli (see Table 1). In the picture-pointing task, participants completed an average of 29.78 trials for words, and 29.94 trials for environmental sounds. There was no significant difference in participant's accuracy in identifying words vs. environmental sounds ($t(17) = .47$, $p = .64$). Further, there was no significant difference in reaction times to identify the visual referent for words vs. environmental sounds ($t(17) = .22$, $p = .82$) (see Table 1).

3.2 Session 2: ERP Results

The overall ERP response for words and environmental sounds was similar in morphology and scalp distribution (Figures 2 & 3). Broadly, the two sound types show a similar pattern of ERP components across the scalp starting with a N100 peaking near 100 ms, followed by a P200 at 175 ms, and a N200 peaking around 250 ms. After the N200 the ERPs are largely characterized by slower and negative-going waves that last through the end of the recording epoch. Further, effects were largest at centro-parietal electrode sites for both sound types (see Figure 2). However, condition-specific differences were present (for a complete description of the results see Tables 2 - 4).

3.2.1 Spoken Words—For words, there was a significant main effect of condition and a significant condition \times ROI interaction in all time windows of interest (200 – 400 ms, 400 – 600 ms, 600 – 800 ms, and 800 – 1000 ms). Linear contrasts revealed a pattern of effects in which between-category violations were significantly more negative than matches throughout the epoch. This effect appeared at central-parietal electrode sites in the early time window (200 – 400 ms) and became more broadly distributed as time unfolded (400 – 1000 ms). Within-category violations were significantly more negative than matches at central-parietal sites in all time windows. Finally, from 600 – 1000 ms words exhibited a graded effect such that between-category violations were significantly more negative than within-category violations and matches, and within-category violations were significantly more negative than matches.

3.2.2 Environmental Sounds—For environmental sounds, there was a significant main effect of condition in the 200 – 400 ms, and 800 – 1000 ms time windows, and no significant condition x ROI interactions. Linear contrasts revealed that between-category violations were significantly more negative than matches in the early time window (200 - 400 ms) and at frontal electrode sites in the late time window (800 – 1000ms). Within-category violations were significantly more negative than matches from 800 – 1000 ms post-stimulus onset.

4. Discussion and Conclusion

The overarching goal of the current study was to determine whether words and environmental sounds are processed and organized similarly within the first two years of life. Prior to participating in an ERP task, participants were subjected to a language and environmental sound familiarization and assessment. The familiarization phase was to ensure that each child had similar prior exposure to each of the 30 concepts tested during the ERP task. The assessment phase, 1.) gauged children's *a priori* familiarity with the word and environmental sound stimuli through parental ratings, and 2.) measured comprehension of the words and sounds with a picture-pointing task as a manipulation check to insure comparable comprehension of word-referent and sound-referent relations prior to the ERP study. Parents rated the environmental sounds less familiar than the words, though the average rating for environmental sounds was 4.95 (out of 7), suggesting that parents were relatively certain their children would be familiar with the environmental sound stimuli. Results from the picture-pointing task revealed that recognition for sound-object associations for words and environmental sounds was quite similar: the speed and accuracy of word and environmental sound identification was statistically indistinguishable. These findings are comparable with results obtained using a visually based paradigm, which showed word and environmental sound recognition (accuracy and speed of processing) to be strikingly similar from 15 to 20 months (Cummings et al., 2009).

Although we find no significant difference in performance on the picture-pointing task for words and environmental sounds, it must be noted that overall performance was lower than 50% (~ 13 correct out of 30) for both sound types. This level of performance for both words and environmental sounds may be due to two factors: 1) the decontextualized nature of the task, and 2.) the partially automated nature of the procedure. First, compared with visually based measures or parent report measures of comprehension, it has been shown that because of the additional demands of executing an action, haptic measures gauge the most robust levels of understanding, and thus, may not capture developing knowledge (Diamond, 1985; Baillargeon, DeVos & Graber, 1989; Hofstadter & Reznick, 1996; Gurteen et al., 2011). Nevertheless, haptic responses have been shown to be a reliable predictor of later language abilities (Friend et al., 2003, 2008; Ring & Fenson, 2000; Woodward et al., 1994). Second, as previously mentioned, the picture-pointing task used in the current study was an adaptation of the well-documented Computerized Comprehension Task. The CCT is a completely experimenter-controlled task. The current task was modified from the traditional CCT paradigm to make it partially automated. That is, in order to insure we were testing comprehension on the exact words and sounds used in the subsequent EEG portion of the study, it was important to automate the procedure such that the target word or sound was produced by the computer instead of the experimenter (the experimenter could not have

produced the environmental sounds verbally). Having the computer instead of the experimenter produce the target words and sounds may have attenuated touches to the screen for both stimulus types. Recall however that the primary goal of this task was to insure that there were no significant differences in children's comprehension of these words and environmental sounds that may drive any observed differences in the subsequent EEG task. Indeed, this task captured the considerable variability in comprehension abilities seen at this age (performance ranged from 17% correct (5 out of 30) to 97% correct (29 out of 20)), but no a priori difference in accuracy for words versus environmental sounds. Arguably, the most sensitive measure of children's knowledge is the ERP procedure since this taps non-volitional neural activity. As we will describe below, these results reveal an impressive sensitivity to category violations that varies with sound class. This is particularly interesting given the absence of a priori differences in comprehension.

After the language and environmental sound assessment, children participated in an ERP task to measure how semantic relatedness effects the processing and organization of words vs. environmental sounds. We varied the degree of semantic violation between an auditory stimulus and a preceding pictorial context for both sound classes with the following two aims: 1.) To determine whether children in their 2nd year show N400 incongruity effects for environmental sounds preceded by pictures constituting between-category violations, and 2.) To examine semantic memory organization of words and environmental sounds at 20 months by evaluating between- and within-category violations. Results showed that the electrophysiological marker of semantic processing (N400) can be observed in young children's ERP response to environmental sounds (Aim 1), however the processing of words and environmental sounds is differentially influenced by the degree of the semantic violation (Aim 2).

For words, both within and between-category violations exhibited a significantly greater negative response compared to matches starting between 200 - 400 ms and lasting throughout all time windows of interest. Further, we found a graded response in the mid-latency time window (600 – 800 ms), such that both within- and between-category violations exhibited significantly greater negative responses compared to matches, with between-category violations exhibiting greater negative amplitudes than within-category violations. This finding is consistent with previous behavioral and ERP results that show in the 2nd year of life, children organize their lexical network based on the associative and featural similarities among concepts to which words refer (Areja-Trejo et al., 2009; 2013; Hendrickson & Sundara, 2016; Plunkett & Styles, 2009, von Koss Torkildson et al, 2006; Willitz, et al., 2013).

Environmental sounds, like words, exhibited a significantly greater negative response to between-category violations compared to matches at 20 months. This effect was present from 200 – 400 ms and at frontal electrode sites from 800 – 1000 ms. These results are in line with previous behavioral evidence, and behavioral results within the current study that demonstrate children from 15-25 months display similar performance in their recognition of familiar words and environmental sounds when presented with between-category pictures (Cummings et al., 2009). Further, these results are consistent with adult research that

reliably shows significant N400 effects to between-category picture-environmental sound violations.

In contrast to words that show a greater negative response to within-category violations compared to matches early and consistently, for environmental sounds at frontal electrode sites, matches were numerically more negative than within-category violations from 200 – 800 ms, however this difference did not reach significance. This pattern changed late in the epoch as within-category violations exhibited a significantly greater negative response compared to matches from 800 – 1000 ms.

The current results with 20-month-olds are similar to results obtained on an analogous ERP task with adults (Hendrickson et al., 2015). For adults, it was found that words exhibited a graded pattern of effects in which both far violations (e.g., bird instead of dog) and near violations (e.g., cat instead of dog) were more negative than matches starting at 300 ms and continuing throughout the epoch. Conversely, for environmental sounds, only between-category violations exhibited greater negative responses compared to matches in the early time window, though a graded pattern similar to that of words was exhibited at a later latency time window (400 – 500 ms).

What accounts for the different pattern of results for words and environmental sounds? One interpretation centers on differences in the early occurring perceptual or phonological processes. Word recognition progresses through a series of phonetic and phonemic processing stages that convert the raw acoustic signal to a lexical unit (Frauenfelder & Tyler 1987; Indefrey & Levelt, 2004). Arguably the predominant interpretation of the functional significance of the N400, as it applies to lexical processing, is that it represents the junction between feed-forward perceptual information (i.e., raw acoustic signal to phonological processing) and a dynamic, multi-modal semantic memory (see Kutas & Federmeir, 2011). Further, in adults and young children, ERPs between 200 – 500 ms post word onset have been shown to index pre-lexical processing that is both linguistically mediated (phonology, familiarity, initial activation of word representation) and domain-general (attention). Specifically, there is evidence of a negative deflection for matches compared to semantic violations in frontal electrode sites at ~200 ms prior to the characteristic semantically-mediated N400, that is modulated by various phonologically related processes (Brunellière, & Soto-Faraco, 2015; Dehaene-Lambertz, & Dehaene, 1994; Friederich & Friederici, 2005; Molfese, Burger-Judisch, & Hans, 1991; Molfese, 1989; Simos & Molfese, 1997; Thierry, Vihmen, & Roberts, 2009).

Therefore, pre-semantic and semantic processes, as indicated by the N400 and other earlier negative components, are expected during cross-modal priming with *words* in the 2nd year, but it is unknown whether environmental sound processing also depends on semantic processes or whether it is largely based on early perceptual or phonological processes. That is, differences between the processing of words versus environmental sounds at this age might not only be quantitative in nature (N400 amplitude modulation), but qualitative (different ERP components). Indeed, previous work with adults demonstrates that environmental sounds elicit an early transient negative response to matches compared to

featurally similar violations at frontal electrode sites (similar results were obtained in the current study but did not reach significance).

Whether these early effects for environmental sounds are related to perceptual or phonological processes is still an open question, however there is previous work that suggests environmental sounds are not phonologically processed. First, there is electrophysiological evidence from adults that people do not linguistically mediate or sub-lexicalize environmental sounds, but instead the raw acoustic signal is directly associated with the semantic representation (Schön, Ystad, Kronland-Martine & Besson, 2010). Relatedly, fMRI research suggests words and environmental sounds access phonological and semantic processes in a different fashion. Word processing has been shown to increase activation in a left posterior superior temporal region associated with phonological processing, whereas the processing of environmental sounds has been shown to increased activation in a right fusiform area associated with nonverbal conceptual and structural object processing (Hocking & Price, 2009). Further, phonetic processing for words is required before the semantic representation is accessed (e.g. Indefrey & Levelt, 2004), whereas for environmental sounds, semantic processing is required before phonological retrieval (e.g. Glaser & Glaser, 1989). Given that phonological processes play a larger role in the online processing of words compared to environmental sounds in adults, the early fronto-centrally distributed negativity for matches compared to within-category violations for environmental sounds may center on the initial low level acoustic analysis of environmental sounds, and not phonologically related processes.

Early right posterior Event-Related Desynchronization (ERD) related to the analysis of the acoustical features has been observed for environmental sounds (Lebrun et al., 2001). Therefore, it is possible that environmental sounds require more specific perceptual processing than words. Indeed, environmental sounds within the same category tend to have a similar sound structure (Ballas, 1993). Because environmental sounds tap into meaning directly, without any early conversion of the raw acoustic signal, the process of disambiguating the relation of sounds to meaning may be protracted, creating a more coarsely organized semantic system especially with regard to within category sounds. This is illustrated by the late occurring negative response to within-category violations compared to matches. Therefore, 20-month-olds may misread the raw acoustic signal of the sound, which results in either delayed access or misidentification of the corresponding concept.

An alternative interpretation for the later occurring semantic difference between the sound types focuses not on early perceptual processing, but on differential activation between words and environmental sounds in terms of spreading activation, in which more fine-grained organization structure leads to more selective activation. The traditional time window of the N400 response represents an early point in semantic processing in which meaning associated with the input is being negotiated due to lexically related competitors (Federmeier & Laszlo, 2009). For words, although highly related lexical items are pre-activated, the brain response indicates that the highly related stimulus is indeed a violation (i.e., it is not the match). Conversely, the coarse-grained organization for environmental sounds leads to greater activation throughout the network. Over-activating related items may

create scenarios in which semantic competition is untenable, resulting in delayed arrival at or misidentification of the appropriate meaning.

This interpretation is supported by the findings in the current study. For environmental sounds, toddlers do not show a significantly greater negative response to within-category violations compared to matches until 800 – 1000 ms post stimulus onset. Thus, it may be the case that toddlers over-activate related items earlier in semantic processing (400 – 800 ms), after which they refine the semantic representation and arrive at the appropriate semantic referent downstream (800 -1000 ms) (Federmeier & Laszlo, 2009). Further evidence for this interpretation comes from behavioral results that show that for adults, environmental sound recognition is more susceptible to interference from semantically related competitors (e.g. cow and horse) than is word recognition (Saygin et al., 2005).

A possible limitation in the current study involves potential *a priori* differences in the level of familiarity and exposure to these words and environmental sounds. That is, children could have more *a priori* experience to the words used in the study, and so could have had more time to semantically organize the words. Although there was a significant difference in parent reported familiarity with the word and environmental sound stimuli, we find this interpretation unlikely for multiple reasons. First, we pretested these materials on a group of adults to ensure that the environmental sounds were highly familiar. Second, each child participated in a language and environmental sound assessment to equate levels of exposure to the stimuli used, and test comprehension of the words and environmental sounds. For this assessment, children first participated in a behavioral familiarization task, in which each concept tested during the ERP task was presented 6 times, 3 times with the associated word and 3 times with the associated environmental sound. Subsequently, each child participated in a picture-pointing task to test their comprehension of the words and environmental sounds. Performance on this task revealed that object recognition and speed of processing was nearly identical for both sound types.

Although it could be argued that the sample size of the current study is small ($N=19$) relative to other types of developmental work in infants, it must be noted that the N400 component is very robust, and can be observed with very few participants (Luck, 2014). Indeed, prior studies using ERPs at this age suggest an N of 16 is sufficient to detect the N400 component (Friedrich & Friederici, 2008; Mani, Mills, & Plunkett, 2012; Mills, Prat, Zangl, Stager, Neville, & Werker, 2004). Importantly, effect sizes within the present study were robust, especially in the later time windows. Thus, the sample size in the present study was appropriate to the research question and method with sufficient power to detect the N400 component with robust effect sizes.

5. Conclusion

The current study provides evidence that the electrophysiological marker of semantic processing (N400) can be observed in young children's ERP response to environmental sounds. However, this response does not differentiate readily between environmental sounds within the same category. Overall results from this study suggest that like adults, the young

brain differentially processes semantic information associated with words and environmental sounds.

Acknowledgments

This research was supported by NIH Training Grant 5T32DC007361, NIH R01HD068458, and R01DC009272. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NICHD, NIDCD, or the NIH.

References

- Aramaki M, Marie C, Kronland-Martinet R, Ystad S, & Besson M (2010). Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. *Journal of Cognitive Neuroscience*, 22(11), 2555–2569. [PubMed: 19929328]
- Arias-Trejo N, & Plunkett K (2009). Lexical–semantic priming effects during infancy. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3633–3647. [PubMed: 19933137]
- Arias-Trejo N, & Plunkett K (2013). What’s in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, 128(2), 214–227. [PubMed: 23688648]
- Ballas J (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 250–267. [PubMed: 8473838]
- Borovsky A, Ellis EM, Evans JL and Elman JL (2015). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science*. doi: 10.1111/desc.12343
- Brunellière A, & Soto-faraco S (2015). The interplay between semantic and phonological constraints during spoken-word comprehension. *Psychophysiology*, 52, 46–58. [PubMed: 25039688]
- Campbell AL, & Namy LL (2003). The role of social-referential context in verbal and nonverbal symbol learning. *Child Development*, 74(2), 549–563. [PubMed: 12705572]
- Chen YC, & Spence C (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1554. [PubMed: 21688942]
- Cummings A, & eponien R (2010). Verbal and nonverbal semantic processing in children with developmental language impairment. *Neuropsychologia*, 48(1), 77–85. [PubMed: 19698728]
- Cummings A, eponien R, Dick F, Saygin AP, & Townsend J (2008). A developmental ERP study of verbal and non-verbal semantic processing. *Brain Research*, 1208, 137. [PubMed: 18387601]
- Cummings A, eponien R, Koyama A, Saygin AP, Townsend J, & Dick F (2006). Auditory semantic networks for words and natural sounds. *Brain Research*, 1115(1), 92–107. [PubMed: 16962567]
- Cummings A, Saygin AP, Bates E, & Dick F (2009). Infants’ recognition of meaningful verbal and nonverbal sounds. *Language Learning and Development*, 5(3), 172–190. [PubMed: 20228882]
- Dehaene-Lambertz G, & Dehaene S (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, 370(6487), 292. [PubMed: 8035876]
- Dale PS, & Fenson L (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28(1), 125–127.
- Daltrozzo J, & Schön D (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience*, 21(10), 1882–1892. [PubMed: 18823240]
- Delle Luche C, Durrant S, Floccia C, & Plunkett K (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science* 17(6), 948–955. [PubMed: 24628995]
- Federmeier KD, & Kutas M (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41(4), 469–495.
- Federmeier KD, McLennan DB, Ochoa E, & Kutas M (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2), 133–146. [PubMed: 12212662]

- Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ, Tomasello M, Mervis CB, & Stiles J (1994). Variability in early communicative development. *Monographs of the society for research in child development*, *i*–185. [PubMed: 8047076]
- Frauenfelder UH, & Tyler LK (1987). The process of spoken word recognition: An introduction. *Cognition*, *25*(1–2), 1–20. [PubMed: 3581722]
- Frey A, Aramaki M, & Besson M (2014). Conceptual priming for realistic auditory scenes and for auditory words. *Brain and Cognition*, *84*(1), 141–152. [PubMed: 24378910]
- Friedrich M, & Friederici AD (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of cognitive neuroscience*, *16*(8), 1465–1477. [PubMed: 15509391]
- Friedrich M, & Friederici AD (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *Neuroreport*, *16*(6), 653–656. [PubMed: 15812327]
- Friedrich M, & Friederici AD (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, *19*(18), 1757–1761. [PubMed: 18955904]
- Friedrich M and Friederici AD (2015). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*. doi: 10.1111/desc.12357
- Friend M, & Keplinger M (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, & Computers*, *35*(2), 302–309.
- Friend M, Schmitt SA, & Simpson AM (2012). Evaluating the predictive validity of the Computerized Comprehension Task: Comprehension predicts production. *Developmental psychology*, *48*(1), 136. [PubMed: 21928878]
- Friend M, & Zesiger P (2011). Une réplique systématique des propriétés psychométriques du Computerized Comprehension Task dans trois langues. *Enfance*, *2011*(03), 329–344.
- Glaser WR, & Glaser MO (1989). Context effects in stroop-like word and picture processing. *Journal of Experimental Psychology: General*, *118*(1), 13. [PubMed: 2522504]
- Gygi B (2001). Factors in the identification of environmental sounds. Unpublished doctoral dissertation, Indiana University–Bloomington.
- Hendrickson K, & Sundara M (2016). Fourteen-month-olds' decontextualized understanding of words for absent objects. *Journal of child language*, 1–16
- Hendrickson K, Walenski M, Friend M, & Love T (2015). The organization of words and environmental sounds in memory. *Neuropsychologia*, *69*, 67–76. [PubMed: 25624059]
- Hirsh-Pasek K, & Golinkoff RM (1996). The intermodal preferential looking paradigm: a window onto emerging language comprehension In McDaniel D, McKee C & Cairns HS (Eds.), *Methods for assessing children's syntax* (pp. 105–124). Cambridge, MA: MIT Press.
- Hocking J, Dzafic I, Kazovsky M, & Copland DA (2013). NESSTI: norms for environmental sound stimuli. *PloS One*, *8*(9), e73382. [PubMed: 24023866]
- Hocking J, & Price CJ (2009). Dissociating verbal and nonverbal audiovisual object processing. *Brain and language*, *108*(2), 89–96. [PubMed: 19101025]
- Holcomb PJ, & Neville HJ (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes*, *5*(4), 281–312.
- Ibáñez A, López V, & Cornejo C (2006). ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. *Brain and Language*, *98*(3), 264–275. [PubMed: 16782185]
- Indefrey P, & Levelt WJ (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101–144. [PubMed: 15037128]
- Kay P (1971). Taxonomy and semantic contrast. *Language*, 866–887.
- Kiefer M (2001). Perceptual and semantic sources of category-specific effects: event-related potentials during picture and word categorization. *Memory & Cognition*, *29*(1), 100–116. [PubMed: 11277454]
- Kutas M, & Federmeier KD (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 621–647.

- Kutas M, & Hillyard SA (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203–205. [PubMed: 7350657]
- Lakens D (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for *t*-tests and ANOVA;s. *Frontiers in Psychology*, 4,863. [PubMed: 24324449]
- Lebrun N, Clochon P, Etevenon P, Lambert J, Baron JC, & Eustache F (2001). An ERD mapping study of the neurocognitive processes involved in the perceptual and semantic analysis of environmental sounds and words. *Cognitive Brain Research*, 11(2), 235–248. [PubMed: 11275485]
- Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, & DeYoe EA (2005). Distinct cortical pathways for processing tool versus animal sounds. *The Journal of Neuroscience*, 25(21), 5148–5158. [PubMed: 15917455]
- Luck SJ (2014). *An introduction to the event-related potential technique*. MIT press.
- Mani N, Mills DL, & Plunkett K (2012). Vowels in early words: an event-related potential study. *Developmental Science*, 15(1), 2–11. [PubMed: 22251287]
- May L, & Werker JF (2014). Can a Click be a Word?: Infants' Learning of Non-Native Words. *Infancy*, 19(3), 281–300.
- Mills DL, Conboy B, & Paton C (2005). Do changes in brain organization reflect shifts in symbolic functioning. *Symbol use and symbolic representation*, 123–153.
- Mills DL, Prat C, Zangl R, Stager CL, Neville HJ, & Werker JF (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, 16(8), 1452–1464. [PubMed: 15509390]
- Molfese DL (1989). Electrophysiological correlates of word meanings in 14-month-old human infants. *Developmental Neuropsychology*, 5(2–3), 79–103.
- Molfese DL, Burger-Judisch LM, & Hans LL (1991). Consonant discrimination by newborn infants: Electrophysiological differences. *Developmental Neuropsychology*, 7(2), 177–195.
- Murphy GL, Hampton JA, & Milovanovic GS (2012). Semantic memory redux: An experimental test of hierarchical category representation. *Journal of Memory and Language*, 67(4), 521–539. [PubMed: 23243336]
- Namy LL, Campbell AL, & Tomasello M (2004). The changing role of iconicity in non-verbal symbol learning: A U-shaped trajectory in the acquisition of arbitrary gestures. *Journal of Cognition and Development*, 5(1), 37–57.
- Namy LL, & Waxman SR (1998). Words and gestures: Infants' interpretations of different forms of symbolic reference. *Child Development*, 69(2), 295–308. [PubMed: 9586206]
- Newman AJ, Tremblay A, Nichols ES, Neville HJ, & Ullman MT (2012). The influence of language proficiency on lexical semantic processing in native and late learners of English. *Journal of Cognitive Neuroscience*, 24(5), 1205–1223. [PubMed: 21981676]
- Nigam A, Hoffman JE, & Simons RF (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience*, 4(1), 15–22. [PubMed: 23967854]
- Orgs G, Lange K, Dombrowski JH, & Heil M (2006). Conceptual priming for environmental sounds and words: an ERP study. *Brain and Cognition*, 62(3), 267–272. [PubMed: 16793186]
- Orgs G, Lange K, Dombrowski JH, & Heil M (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience Letters*, 436(2), 133–137. [PubMed: 18378076]
- Özcan E, & Egmond RV (2009). The effect of visual context on the identification of ambiguous environmental sounds. *Acta Psychologica*, 131(2), 110–119. [PubMed: 19403108]
- Pace A, Carver L, & Friend M (2013). Event-related potentials to intact and disrupted actions in children and adults. *Journal of Experimental Child Psychology*, <http://dx.doi.org/10.1016/j.jecp.2012.10.013>
- Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, & Di Russo F (2005). Separate neural systems for processing action- or non-action-related sounds. *Neuroimage*, 24(3), 852–861. [PubMed: 15652320]
- Plante E, Petten CV, & Senkfor AJ (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia*, 38(13), 1669–1684. [PubMed: 11099725]

- Poulin-Dubois D, Bialystok E, Blaye A, Polonia A, & Yott J (2013). Lexical access and vocabulary development in very young bilinguals. *International Journal of Bilingualism*, 17(1), 57–70. [PubMed: 24761135]
- Rämä P, Sirri L, & Serres J (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18- and 24-month old children. *Brain and language*, 125(1), 1–10. [PubMed: 23435193]
- Reid VM, Hoehl S, Grigutsch M, Groendahl A, Parise E, & Striano T (2009). The neural correlates of infant and adult goal prediction: evidence for semantic processing systems. *Developmental psychology*, 45(3), 620. [PubMed: 19413420]
- Rosch E, Mervis CB, Gray WD, Johnson DM, & Boyes-Braem P (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439.
- Sajin SM, & Connine CM (2014). Semantic richness: The role of semantic features in processing spoken words. *Journal of Memory and Language*, 70, 13–35
- Saygin AP, Dick F, & Bates E (2005). An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. *Behavior Research Methods*, 37(1), 99–110. [PubMed: 16097349]
- Schirmer A, Soh YH, Penney TB, & Wyse L (2011). Perceptual and conceptual priming of environmental sounds. *Journal of Cognitive Neuroscience*, 23(11), 3241–3253. [PubMed: 21281092]
- Schneider TR, Engel AK, & Debener S (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, 55(2), 121–132.
- Schön D, Ystad S, Kronland-Martinet R, & Besson M (2010). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience*, 22(5), 1026–1035. [PubMed: 19583472]
- Sheehan EA, Namy LL, & Mills DL (2007). Developmental changes in neural activity to familiar words and gestures. *Brain and Language*, 101(3), 246–259. [PubMed: 17250885]
- Simos PG, & Molfese DL (1997). Event-related potentials in a two-choice task involving within-form comparisons of pictures and words. *International journal of neuroscience*, 90(3–4), 233–253. [PubMed: 9352430]
- Styles SJ, & Plunkett K (2009). How do infants build a semantic system? *Language and Cognition*, 1(1), 1–24.
- Thierry G, Vihman M, & Roberts M (2003). Familiar words capture the attention of 11-month-olds in less than 250 ms. *Neuroreport*, 14(18), 2307–2310. [PubMed: 14663181]
- Van Petten C, & Rieffers H (1995). Conceptual relationships between spoken words and ES: Event-related brain potential measures. *Neuropsychologia*, 33(4), 485–508. [PubMed: 7617157]
- von Koss Torkildsen J, Sannerud T, Syversen G, Thormodsén R, Simonsen HG, Moen I, Smith L, & Lindgren M (2006). Semantic organization of basic-level words in 20-month-olds: An ERP study. *Journal of Neurolinguistics*, 19(6), 431–454.
- von Koss Torkildsen J, Svangstu JM, Hansen HF, Smith L, Simonsen HG, Moen I, & Lindgren M (2008). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience*, 20(7), 1266–1282. [PubMed: 18284350]
- Vouloumanos A, & Werker JF (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental science*, 10(2), 159–164. [PubMed: 17286838]
- Willits JA, Wojcik EH, Seidenberg MS, & Saffran JR (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: evidence from auditory priming. *Infancy*, 18(6), 1053–1075.

Research Highlights

- Recorded ERPs as 20-month-olds viewed images while hearing words or environmental sounds that matched the image or exhibited a within- or between-category semantic violation.
- Electrophysiological marker of semantic processing (N400) was observed in response to between-category violations for both words and environmental sounds.
- For words, within-category violations demonstrated greater negative response than matches, however ERP response did not differentiate readily between environmental sounds that were featurally similar.
- More consistent fine-grained organizational structure apparent for words than sounds.

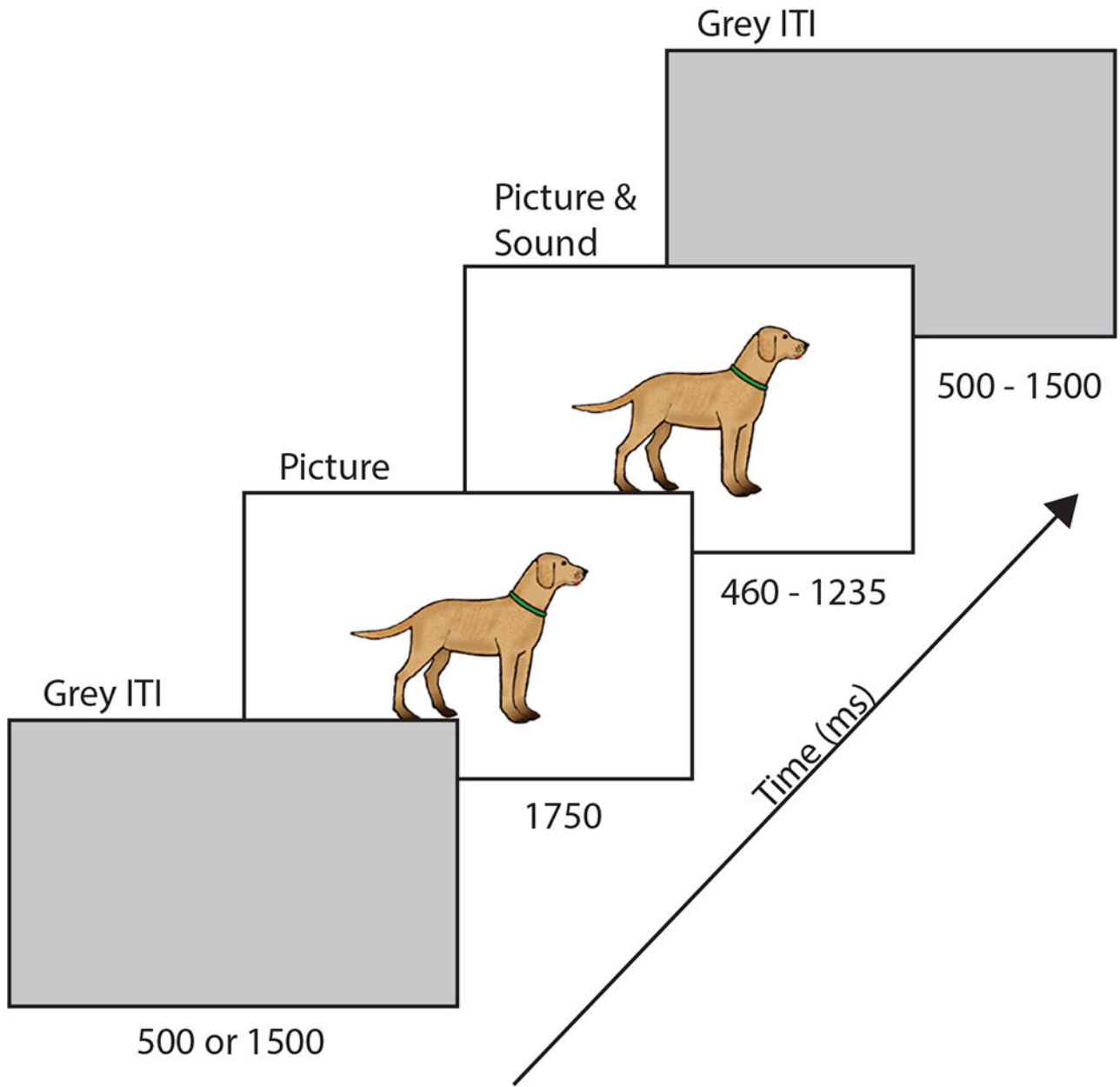


Figure 1.

Schematic of a single trial. For each sound type (word, environmental), participants were presented with a colorful line drawing of a familiar concept for 1750 ms before hearing a sound (duration 460 – 1235 ms) from one of three conditions (match, within-category violation, between-category violation). Pictures disappeared at the offset of the sound. A variable inter-trial interval grey screen (duration 500 or 1500 ms) was presented at the offset of the picture and sound.

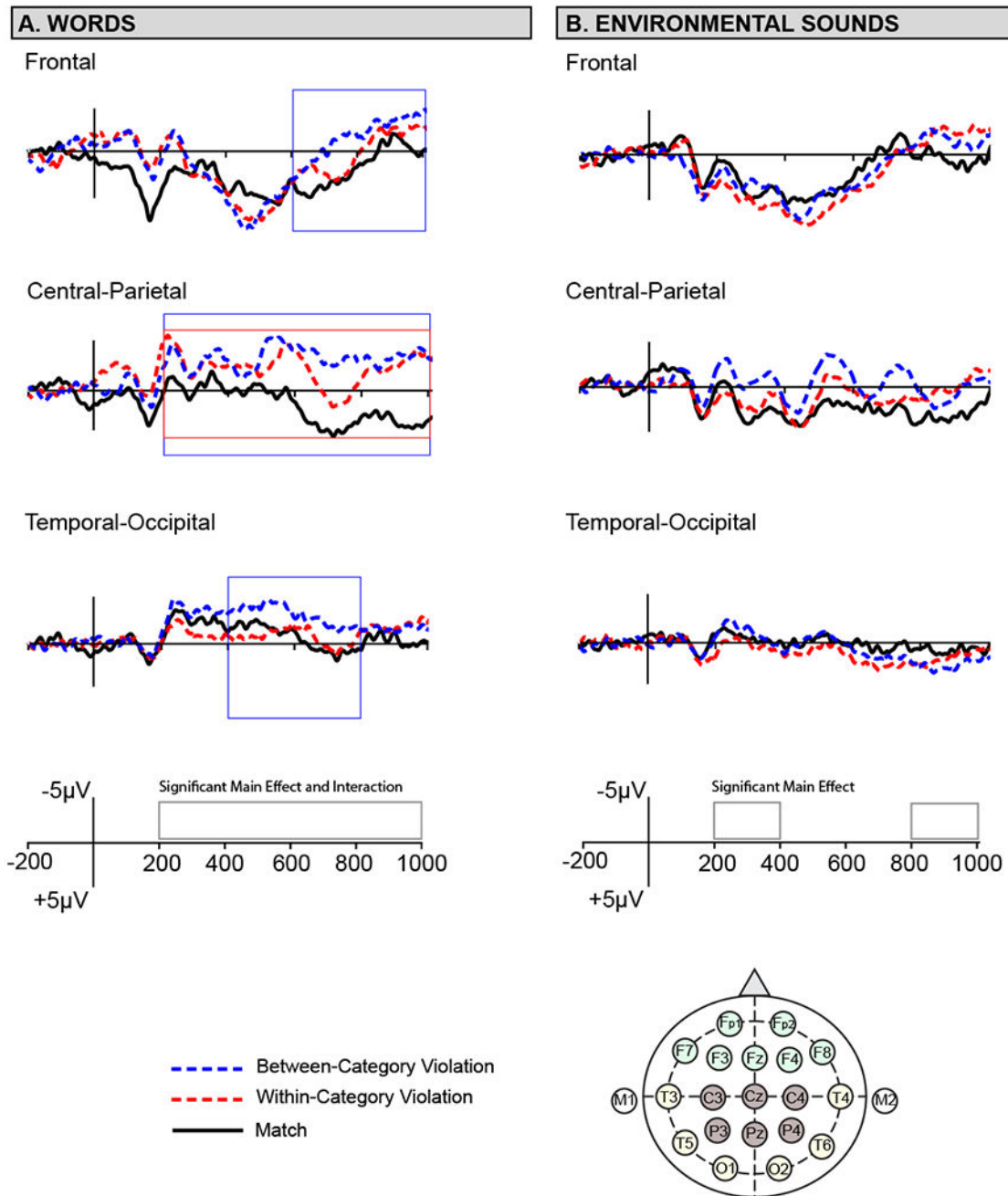


Figure 2. Grand average ERP waveforms for the 3 conditions (match, within-category violation, between category violation) at 3 ROIs of interest for words (left) and environmental sounds (right).

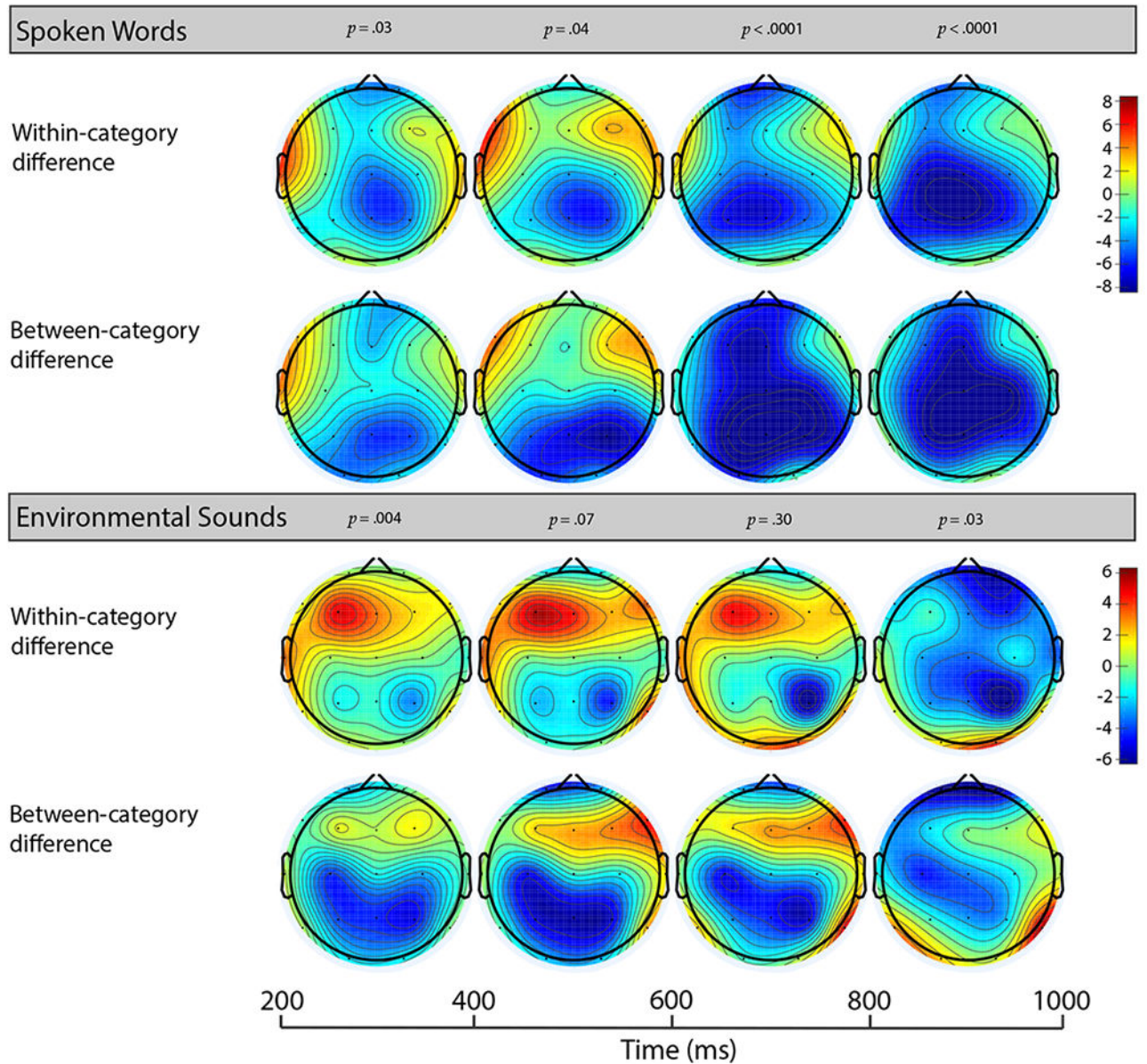


Figure 3.

Voltage maps show average voltage difference (measured as violation – match) for within- and between category violations for words (upper) and environmental sounds (lower) at all 19 electrodes. Mean amplitude voltage calculated between two fixed latencies.

Table 1.

Means and standard errors of parental ratings of word and environmental sound familiarity and performance measures on picture pointing task (accuracy and reaction time).

Sound Type	Parental Rating	Picture Pointing Accuracy	Average Reaction Time
Word	5.62 (.18)	13.67 (1.31)	2426.92 (168.75)
Environmental	4.95 (.14)	12.94 (1.04)	2378.73 (226.52)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.**Omnibus results.**

Type III *F*-tests for the main effects and interactions of Condition (match, within-category, between-category), Anteriority (frontal, central, parietal), Laterality (left, center, right), and their interactions, within the four time windows of interest.

Coefficients	200 - 400 ms	400 - 600 ms	600 - 800 ms	800 - 1000 ms
Words				
Condition	3.6 [*]	3.23 [*]	28.8 ^{***}	21.11 ^{***}
ROI	37.2 ^{***}	102.6 ^{***}	9.6 ^{***}	1.54
Condition x ROI	2.33 [*]	4.28 ^{**}	2.82 [*]	4.14 ^{**}
Environmental Sounds				
Condition	5.65 ^{**}	2.65	1.2	3.5 [*]
ROI	29.85 ^{***}	31.88 ^{***}	1.04	10.81 ^{***}
Condition x ROI	1.68	1.92	2.14	2.18

*
p < .05,

**
p < .01,

p < .001

Table 3.

Post-hoc linear contrasts of Condition.

Coefficients	200 - 400 ms			400 - 600 ms			600 - 800 ms			800 - 1000 ms						
	B (SE)	t-value	CI	<i>g_{av}</i>	B (SE)	t-value	CI	<i>g_{av}</i>	B (SE)	t-value	CI	<i>g_{av}</i>				
Words																
Between - Within	-0.89 (0.66)	1.35	(-2.19, 0.41)	-0.15	-1.42 (0.74)	1.92	(-2.87, 0.03)	-0.26	-2.77 (0.71)	3.9***	(-4.16, -1.38)	-0.44	-0.72 (0.72)	1.01	(-2.11, 0.68)	-0.11
Between - Match	-1.77 (0.66)	2.68**	(-3.08, -0.48)	-0.27	-1.77 (0.74)	2.4*	(-3.23, -0.32)	-0.25	-5.38 (0.71)	7.59***	(-6.78, -3.99)	-0.83	-4.31 (0.72)	6.06***	(-5.7, -2.91)	-0.57
Within - Match	-0.89 (0.66)	1.34	(-2.19, 0.41)	-0.15	-0.35 (0.74)	0.48	(1.80, -1.20)	-0.04	-2.61 (0.71)	3.68***	(-4.01, -1.22)	-0.40	-3.59 (0.72)	5.06***	(-4.90, -2.20)	-0.52
Environmental Sounds																
Between - Within	-2.12 (0.64)	3.3**	(-3.38, -0.86)	-0.51									1.02 (0.80)	1.26	(-0.56, 2.59)	0.24
Between - Match	-1.42 (0.64)	2.21*	(-2.68, -0.16)	-0.26									-1.11 (0.80)	1.38	(-2.69, -0.47)	-0.25
Within - Match	0.70 (0.64)	1.09	(-0.56, 1.96)	0.14									-2.13 (0.80)	2.64**	(-3.70, -0.55)	-0.42

* p < .05,
 ** p < .01,
 *** p < .001

Table 4.

Post-hoc linear contrasts of significant Condition x ROI interactions.

	Frontal					Central-Parietal					Temporal-Occipital						
	B (SE)	t-value	CI	β_{op}	B (SE)	t-value	CI	β_{op}	B (SE)	t-value	CI	β_{op}	B (SE)	t-value	CI	β_{op}	
200 - 400 ms	Between - Within	-0.94 (1.09)	0.09	(-2.2, 2.04)	-0.01	0.07 (1.17)	0.06	(-2.24, 2.38)	0.01	-2.66 (1.17)	2.26*	(-4.96, -0.35)	-0.49	-2.66 (1.17)	2.26*	(-4.96, -0.35)	-0.49
	Between - Match	-0.94 (1.09)	.87	(-3.09, 1.19)	-0.12	-3.21 (1.17)	2.73**	(-5.52, -.91)	-0.34	-1.18 (1.17)	1	(-3.49, 1.13)	-0.19	-1.18 (1.17)	1	(-3.49, 1.13)	-0.19
	Within - Match	0.85 (1.09)	0.78	(-1.99, 1.99)	-0.12	-3.29 (1.17)	2.79**	(-5.59, -0.98)	-0.37	1.47 (1.17)	1.25	(-0.84, 3.78)	0.26	1.47 (1.17)	1.25	(-0.84, 3.78)	0.26
400 - 600 ms	Between - Within	1.00 (1.22)	0.08	(-2.28, 2.48)	0.01	-0.86 (1.31)	.65	(-3.43, 1.72)	-0.12	-3.51 (1.31)	2.67**	(-6.08, -0.93)	-0.56	-3.51 (1.31)	2.67**	(-6.08, -0.93)	-0.56
	Between - Match	1.65 (1.22)	1.37	(-0.73, 4.04)	0.20	-4.37 (1.31)	3.33***	(-6.95, -1.79)	-0.51	-2.61 (1.31)	1.99*	(-5.19, -0.04)	-0.42	-2.61 (1.31)	1.99*	(-5.19, -0.04)	-0.42
	Within - Match	1.56 (1.22)	1.28	(-0.82, 3.95)	0.20	-3.51 (1.31)	2.68**	(-6.10, -0.94)	-0.41	0.90 (1.31)	0.68	(-1.68, 3.47)	0.14	0.90 (1.31)	0.68	(-1.68, 3.47)	0.14
600 - 800 ms	Between - Within	-2.58 (1.17)	2.21*	(-4.87, -0.29)	-0.30	-3.61 (1.26)	2.87**	(-6.08, -1.14)	-0.48	-2.12 (1.26)	1.68	(-4.59, 0.35)	-0.36	-2.12 (1.26)	1.68	(-4.59, 0.35)	-0.36
	Between - Match	-4.29 (1.17)	3.68***	(-6.57, -2.00)	-0.43	-8.63 (1.26)	6.86***	(-11.11, -6.16)	-1.17	-3.23 (1.26)	2.56*	(-5.70, -0.76)	-0.63	-3.23 (1.26)	2.56*	(-5.70, -0.76)	-0.63
	Within - Match	-1.71 (1.17)	1.46	(-3.99, 0.58)	-0.18	-5.03 (1.26)	3.99***	(-7.50, -2.56)	-0.55	-1.12 (1.26)	0.88	(-3.58, 1.36)	-0.26	-1.12 (1.26)	0.88	(-3.58, 1.36)	-0.26
800 - 1000 ms	Between - Within	-1.45 (1.17)	1.24	(-3.74, 0.84)	-0.16	-0.97 (1.26)	0.77	(-3.45, 1.50)	-0.10	0.28 (1.26)	0.22	(-2.20, 2.75)	0.05	0.28 (1.26)	0.22	(-2.20, 2.75)	0.05
	Between - Match	-3.56 (1.17)	3.05**	(-5.85, -1.27)	-0.34	-7.87 (1.26)	6.23***	(-10.34, -5.39)	-0.85	-1.50 (1.26)	1.19	(-3.98, 0.97)	-0.26	-1.50 (1.26)	1.19	(-3.98, 0.97)	-0.26
	Within - Match	-2.11 (1.17)	1.81	(-4.40, 0.18)	-0.23	-6.89 (1.26)	5.46***	(-9.37, -4.42)	-0.71	-1.77 (1.26)	1.41	(-4.25, 0.70)	-0.36	-1.77 (1.26)	1.41	(-4.25, 0.70)	-0.36

* p < .05,
 ** p < .01,
 *** p < .001