HHS Public Access

The organization of words and environmental sounds in memory[☆]

Kristi Hendrickson[a,c,d,*], **Matthew Walenski**[c], **Margaret Friend**[b], and **Tracy Love**[a,c]

[a]Center for Research in Language, University of California, San Diego, USA

[b]Department of Psychology, San Diego State University, USA

[c]School of Speech, Language, and Hearing Sciences, San Diego State University, USA

[d]Joint Doctoral Program in Language and Communicative Disorders, San Diego State University, USA

Abstract

In the present study we used event-related potentials to compare the organization of linguistic and meaningful nonlinguistic sounds in memory. We examined N400 amplitudes as adults viewed pictures presented with words or environmental sounds that matched the picture (Match), that shared semantic features with the expected match (Near Violation), and that shared relatively few semantic features with the expected match (Far Violation). Words demonstrated incremental N400 amplitudes based on featural similarity from 300–700 ms, such that both Near and Far Violations exhibited significant N400 effects, however Far Violations exhibited greater N400 effects than Near Violations. For environmental sounds, Far Violations but not Near Violations elicited significant N400 effects, in both early (300–400 ms) and late (500–700 ms) time windows, though a graded pattern similar to that of words was seen in the midlatency time window (400–500 ms). These results indicate that 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.

Keywords

ERP; N400; Word; Environmental sound; Semantic memory

Uncited references

Griffith et al. (1999), Griffiths et al., (2004), Perrin and Garcýa- Larrea (2003), and Pizzamiglio et al. (2005).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/jfineuropsychologia. 2015.01.035.

1. Introduction

Our ability to interpret the world around us crucially depends on how the brain organizes meaningful auditory information in memory. The organization of semantic memory for one form of meaningful information, linguistic items (e.g. words), has been well investigated, and is based on several factors. Among the most important is featural similarity (i.e. the perceived likeness between concepts), which aids in categorization (Kay, 1971; Murphy et al., 2012; Paczynski and Kuperberg, 2012; Rosch et al., 1976; Sajin and Connine, 2014). Far less is known about how the brain processes and organizes meaningful auditory information that is not linguistic (e.g. environmental sounds). The current paper examines whether semantic information is organized similarly in memory for words and environmental sounds, and specifically whether featural similarity is useful for the organization of environmental sounds in memory. Uncovering how the brain organizes meaning associated with diverse forms of referential auditory information is vital for understanding the relation between language and cognition.

1.1. The processing of words and environmental sounds

Words and environmental sounds share many spectral and temporal characteristics (Gygi, 2001; Shafiro and Gygi, 2004) and are modulated by contextual cues (Ballas and Howard, 1987), item familiarity (Ballas, 1993; Cycowicz and Friedman, 1998), and frequency of occurrence (Ballas, 1993; Cycowicz and Friedman, 1998). Like words, environmental sounds carry deep semantic associations with a corresponding referent (Ballas, 1993). Multiple lines of evidence suggest that words and environmental sounds are processed similarly. It has been shown with behavioral measures (accuracy, response time) that semantically congruent words or pictures can prime environmental sounds, and it has likewise been shown that environmental sounds can prime words or pictures (Ballas, 1993; Chen and Spence, 2011; Özcan and Egmond, 2009; Schneide et al., 2008; Stuart and Jones, 1995). Electrophysiological measures reveal a similar effect. N400 (described in detail below) priming effects (attenuated N400 amplitudes to semantically related compared to unrelated primes) have been found for word or picture targets primed by environmental sounds (Schön et al., 2010; Daltrozzo and Schön, 2009; Frey et al., 2014; Koelsch et al., 2004; Van Petten and Rheinfelder, 1995) and for environmental sound targets primed by words, pictures, or other environmental sounds (Aramak et al., 2010; Cummings et al., 2006, 2008; Cummings and Èeponiene, 2010; Daltrozzo and Schön, 2009; Orgs et al., 2008; Orgs et al., 2006; Plante et al., 2000, Van Petten and Senkfor, 2000; Schirmer et al., 2011; Schön et al., 2010; Van Petten and Rheinfelder, 1995). Indeed several studies of N400 priming effects using bimodal (visual/auditory) stimulus presentation have found similar scalp distributions for the N400 priming effects to words and environmental sounds across multiple ages (Cummings et al., 2006, 2008; Cummings and Èeponiene, 2010; Orgs et al., 2007) Finally, functional imaging results have shown activation to both word and environmental sound stimuli in areas commonly thought of as language specific: left inferior frontal and superior temporal regions (Binder et al., 2000; Leech and Saygin, 2011; Price et al., 2005; Thierry et al., 2003; Tranel et al., 2003) and similar neural networks being implicated in the semantic processing of speech and musical sounds (Koelsch, 2005; Koelsch et al., 2004; Steinbeis and Koelsch, 2008). Despite these similarities, there are some

important differences between words and environmental sounds. These differences exist on multiple dimensions since environmental sounds are non-linguistic. Whereas words have an arbitrary linkage to the items to which they refer, environmental sounds obtain meaning through the causal relation with the event or object that produces them (Ballas and Howard, 1987). Thus, the "lexicon" of environmental sounds is rather small, and tends to converge on a limited number of referents (Ballas, 1993).

Consequently, there is also some empirical support for the notion that distinct mechanisms underlie the processing of each sound type. Behavioral evidence suggests that environmental sound recognition is more susceptible to interference from semantically related competitors (e.g. cow and horse) than word recognition is (Saygin et al., 2005). Additionally, there is evidence from dichotic listening studies that environmental sounds are processed more efficiently in the right hemisphere, whereas words are processed more efficiently in the left hemisphere (Knox and Kimura, 1970; Kimura, 2011). While seemingly at odds with the dichotic listening research, ERP studies (using uni-modal auditory presentation) have found words and environmental sounds exhibit different scalp distributions for N400 priming effects: words showing a larger effect over the right hemisphere, whereas environmental sounds show a larger effect over the left hemisphere (Van Petten and Rheinfelder, 1995; but see above). We note here that the scalp topography of an ERP component does not correspond in any straightforward way to the location of its underlying neural generators, but reflects the summed activity of all generators, which vary in location, strength, and orientation with respect to the scalp. Therefore a right-sided asymmetry at the scalp does not implicate right-hemisphere generation, and results from dichotic listening tasks and ERP scalp topography are not necessarily at odds.

To further bolster the idea that words and environmental sounds indeed call upon different processing routines, and hence different neural networks, functional imaging research has revealed differential intra- and inter- hemispheric activation patterns for words (left angular gyrus, and left anterior and posterior temporal areas) and for environmental sounds (left superior and middle temporal gyri and right superior temporal cortex) (Noppeney et al., 2008; Thierry et al., 2003; Humphries et al., 2001). Finally, using electrophysiological and hemodynamic measures concurrently, Renvall et al. (2012) found that adding background noise affected the recognizability, timing, and location of cortical responses differently for each sound class.

1.2. The effect of featural similarity on N400 responses

All stimuli with referential meaning, whether auditory, visual, orthographic, or pictorial, elicit an N400 component, which is a negative voltage deflection peaking approximately 400 ms post-stimulus onset (Kutas and Federmeier 2011; Kutas and Hillyard, 1980; Kutas and Hillyard, 1983). The prototypical (visual) N400 semantic incongruity effect – the relative amplitude of the waveform compared to another experimental condition (e.g. unprimed target minus primed target) – is typically maximal over right parietal, posterior temporal, and occipital sites. However auditory N400s tend to begin earlier, last longer, and have a somewhat more frontal and less right-biased scalp distribution than visual N400s (Holcomb and Neville,1990; and reviewed in Kutas and Van Petten, 1994). It has been shown that

N400 amplitude (to visual or auditory stimuli) is sensitive to category membership. For instance, following a series of prime words from the same taxonomic category (e.g. flower), N400 amplitudes are larger for target words that belonged to a different category (e.g. apple) than target words belonging to the primed category (e.g. tulip) (Polich, 1985; review in Kutas and Van Petten, 1988).

N400 amplitude for words is not only sensitive to gross category membership (member vs. non-member) but is also incrementally sensitive to differences in featural similarity (i.e. the perceived likeness between concepts as measured by the degree of overlap in their semantic features) (Federmeier and Kutas, 1999, 2002; Federmeier et al., 2002; Ibáñez et al., 2006; Torkildson et al., 2006). We know that the brain often represents feature information in a structured fashion such that neurons responding to similar features are physically close to one another (Hubel and Wiesel, 1972; Tanaka, 1996). If we characterize neural representations of words as a collection of features, then two words that share many features will show similarities in their underlying neural activity (Amuntz and Zilles, 2012; Federmeir et al., 1999b). Indeed an incremental or graded effect in N400 amplitude for words based on featural similarity was first found by investigating the effects of sentential context on semantic memory organization (Federmeier and Kutas 1999; Federmeier et al., 2002). Federmeir and colleagues defined feature likeness in terms of taxonomic semantic categories (e.g. bears and pandas are within the same taxonomic category and therefore share more features than do bears and zebras). Participants were visually presented with sentences that ended in three types of words: expected exemplars (e.g. panda), unexpected exemplars from the same category (e.g. bear), and unexpected exemplars from a different category (e.g. zebra). Both within- and between-category violations exhibited significant N400 effects; however between-category violations (e.g. 'zebra' instead of 'panda') exhibited greater N400 amplitudes than within-category violations (e.g. 'bear' instead of 'zebra') (Federmeier and Kutas, 1999). What's more, the graded effect in N400 amplitudes based on featural similarity has been replicated in visual-auditory priming paradigms across the lifespan (Federmeier et al., 2002; Ibáñez et al., 2006; Torkildson et al., 2006).

Active listening paradigms (i.e. paradigms in which participants are given a concurrent behavioral task to maintain attention), like those mentioned just above, are not mandatory for eliciting the N400 priming effect; the effect has been repeatedly found during attentional blink tasks in which a visual stimuli is not detected due to rapid presentation (Luck et al., 1996; Maki et al., 1997; Vogel et al., 1998), and has even been found for participants who were presented with stimuli while asleep (Ibáñez et al., 2006). Thus active listening does not appear necessary to elicit N400 priming effects, and effects found with passive listening paradigms appear to be comparable to those found during active listening paradigms.

Together these works demonstrate that featural similarities between concepts in the world influence the neural organization of lexical items, and further, that the N400 is sensitive to the organization of lexical-semantic categories in memory. Organizing representations of words based on subtle differences in feature overlap creates a processing benefit for items that are related; transitions between the pattern of activation corresponding to one word and that corresponding to a different but related word are likely easier (Federmeir et al., 1999b). However, no previous evidence we are aware of addresses whether featural similarity

between concepts acts on the organization and subsequent processing of environmental sounds similarly as for words. Thus this is a crucial question for the current paper to address.

1.3. Current study

In this study, we examine the effect of featural similarity on the processing of words and environmental sounds by using a cross-modal sound-picture match/mismatch ERP paradigm, in which we varied feature likeness among within-category pictures and auditory stimuli. Participants viewed pictures (e.g. dog) presented with words or environmental sounds at three levels of featural similarity: those that match the picture (Match: e.g. "dog" or *barking*), those that share semantic features with the expected match (Near Violations: e.g. "cat" or *meowing*), and those that share few semantic features with the expected match (Far Violations: e.g. "lion" or *roaring*).

To elucidate similarities and differences in the structure of words and environmental sounds in memory, our primary interest is how the relative N400 amplitude of Near Violations, a condition not examined in previous studies, compares to the Match condition and the Far Violations for each sound type. For words, we expect to replicate findings that show a graded organization: both Near and Far Violations will exhibit significant N400 effects, however Far Violations will exhibit greater N400 effects than Near Violations.

For environmental sounds we expect Far Violations to exhibit significant N400 effects, consistent with previous work. However, predictions for the Near Violations are not as certain. One possibility, based on behavioral, ERP, and FMRI evidence of similar processing for words and environmental sounds, is that words and environmental sounds are similar with respect to their organization in memory. On this view, environmental sounds should show a similar graded pattern as words, with both Near Violations and Far Violations exhibiting significant N400 effects, and Far Violations exhibiting larger N400 effects than Near Violations. This outcome would suggest that featural similarities between concepts influence the neurocognitive organization of environmental sounds in a graded fashion, similar to that of words.

An alternate possibility, consistent with fMRI findings that pinpoint distinct neural activation for words vs. environmental sounds and behavioral findings that people are less accurate at differentiating within-category members for environmental sounds than words (Saygin et al., 2005), is that words and environmental sounds will not show similar effects of featural similarity. Thus the processing of environmental sounds may not be dependent upon a strict feature match between sounds and concepts. In this case, environmental sounds, unlike words, will exhibit a non-graded effect: Far Violations will show a significant N400 effect but Near Violations will not.

2. Materials and methods

2.1. Participants

This study is part of a larger project examining the neural response to words vs. environmental sounds in infants, toddlers, and adults. Here, we report data for 22 adults (15 women, 7 men, 18–38 years of age, mean age 24). All participants were monolingual native

English speakers and right-handed. Participants were recruited from the undergraduate and graduate populations at San Diego State University and either volunteered to participate or participated for course credit.

2.2. Stimuli

Stimuli for the cross-modal picture-word match/mismatch paradigm were line drawings, auditory words, and environmental sounds of 44 highly familiar concepts such as "dog" (all of which were nouns). A female native English speaker produced the word stimuli (mean duration = 876 ms, SD = 197 ms), which were recorded in a single session in a sound-attenuating booth (sampling at 44.1 Hz, in 16-bit stereo). 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 et al., 2013). Environmental sounds were standardized for sound quality (44.1 kHz, 16 bit, stereo). The duration of words (876 ms, SD = 197 ms) and environmental sounds (1045 ms, SD: 616 ms) did not significantly differ ($t(86)$ = 1.73, p = 0.09). Visual stimuli were black and white line drawings (600×600 pixels) depicting typical exemplars of each concept. The images were digitally edited to remove backgrounds and distracting features.

In order to ensure that the concepts were associated with easily identifiable environmental sounds for the match condition, a Likert scale pretest was conducted. Ten native English-speaking SDSU undergraduates (independent of those who participated in the EEG portion of the study) were presented with 51 images of prototypical members of highly familiar concepts (e.g., dog) paired with an associated environmental sound (e.g., barking). Each image was presented twice, though in a randomized order, each time with a different exemplar of the associated environmental sound. Therefore, 102 presentations of image/sound pairs were presented one at a time with participants asked to rate, on a 1–5 scale (1 = not related and 5 = very/highly related), how well the picture and sound went together. Only those sounds that received a mean rating of 3.5 or higher were included as stimuli (7 of the 51 concepts were excluded because they did not receive a score of 3.5 or higher for either example sound, resulting in a final set of 44 concepts). 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 (see Appendix).

To create stimulus sets based on feature similarity we used similarity scores from existing semantic feature production norms (McRae et al., 2005). McCrae and colleagues asked adults to list different types of features, such as physical (perceptual) properties (how it looks, sounds, smells, feels, and tastes), functional properties (what it is used for and where and when it is used), and other facts (such as where it is from). Concept similarity scores were derived by calculating the cosine between each pair of concepts (on the basis of feature production frequencies). The scores ranged from 1, indicating perfect correspondence (i.e. the cosine between a concept and itself) to 0 indicating complete concept independence (i.e. the concepts have no features in common). To be considered a violation, concept similarity scores had to be between 0.20 and 0.70 for Near Violations and 0.02–0.20 for Far Violations. Moreover, the difference between Near and Far Violations scores for the same

concept had to be greater than 0.3. For both the word and environmental sound conditions, pictures of each concept were paired with one of three types of auditory stimuli: Matches (average concept similarity score = 1); Near Violations (average concept similarity score = 0.4); and Far Violations (average concept similarity score = 0.1). Six of our 44 stimuli (bunny, bug, bee, bird, monkey, and fish) were not included in the McRae norms, so we used scores from the closest prototypical member of that category for which similarity score were available (rabbit, ant, wasp, robin, gorilla, and gold fish, respectively).

Near and Far Violation conditions were created to be only within category: animals, vehicle, tools and household/outside objects. We used solely within-category stimuli for two reasons: firstly, between and within category items can differ not only based on featural similarity, but also by broader animacy distinctions. For example, it has been shown that animate (e.g. dog) and inanimate (e.g. pen) environmental sounds activate distinct neural substrates (Lewis et al., 2005). Thus, differences between violation conditions could be due to the activation of different underlying neural networks and not to changes in feature likeness per se. Secondly, by using only within category members we were able to control for feature likeness in a more systematic way than if we used between category violations, which often times resulted in a similarity score of zero in the McRae et al., 2005 semantic feature production norms.

The same 44 items were used to make Match, Near Violation, and Far Violation featural similarity conditions (Appendix A). Thus the three conditions were largely identical (e.g., *Ringing* is an environmental sound present in all three conditions – Match, Near, Far – as is the corresponding word "telephone"). Note, however, while all 44 items were used in the Match condition, not all 44 items had suitable scores for the Near and Far violation conditions, per the McRae et al., norms. As a result, a portion of the 44 items were used as Near Violations (32 unique pictures) or Far Violations (30 unique pictures) more than once (but no more than 3 times) while others were not included at all (see Appendix for list of stimuli). Therefore the final stimulus list consisted of 44 Match trials, 44 Near Violation trials, and 41 Far Violation trials. Nevertheless, as the three featural similarity conditions for both words and environmental sounds consisted almost entirely of the same stimuli, the conditions were very well controlled for word frequency, imageability, concreteness, phonology, and other properties of the stimuli.

2.3. Design

We used a 2 (sound class: Word, Environmental Sound)×3 (feature similarity: Match, Near and Far violation) within subjects design; sound class was presented in a blocked fashion, so that we had two back-to-back runs that were conducted in a single experimental session. Each of the runs was composed of a single presentation list with 129 trials: 44 in the Match condition, 44 in the Near Violation condition, and 41 in the Far violation condition. Repeated pictures were always at least 10 trials apart in the presentation list. Two versions of each run list for each sound condition were created, the second of each in reverse order of the first.

2.4. Procedure

Participants were seated in a comfortable chair 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 12 min. The only differences between the runs were the type of sound (Word or Environmental) and the trial order (reversed). The order of the runs was counterbalanced such that half the participants received the Word run first. On each trial participants were presented with a line 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° on average). After 1500 ms participants heard a sound from one of three levels of featural similarity (Match, Near Violation, Far Violation). As mentioned above, the current study is part of a larger project designed to include infants and toddlers as well as adults. As infants and toddlers require longer ISIs than adults (Richards, 2003), the longer delay between the onset of the picture and the onset of the paired word or sound was used to maintain consistency between the adult and child versions of this experiment for the larger project. The picture disappeared at the offset of the sound (300–2800 ms). A brief inter-stimulus interval grey screen was presented for 2000 ms, followed by a centrally located fixation cross for 500 ms (Fig. 1). After every 31 trials participants received a 12 s break, and a brief break was given between each run. Presentation of the Matches, Near Violations, and Far Violations was pseudo-randomized across the presentation list such that the same trial type did not appear for more than three consecutive trials.

Participants were asked to maintain their gaze toward the center of the screen and refrain from blinking during the picture presentation, but were not asked to make an overt response to the stimulus. We chose a passive listening paradigm instead of a paradigm requiring participants to respond to each stimulus for three reasons. Firstly, the infant and toddler populations we are testing for the larger project (16- and 24-months respectively) cannot execute reliable behavioral responses during the ERP experiment. Thus to maintain consistency between the adult and child samples we felt it necessary to make the ERP experiment a passive listening task. Secondly, we wanted to have the N400 as clean as possible, without obstruction by motor ERPs, which could differ significantly between conditions; there is evidence that people process environmental sounds faster than words (Cummings et al., 2006, Orgs et al., 2006), thus reaction times could have been faster for environmental sounds compared to words. Thirdly, N400 priming effects are reliably found during passive listening tasks (Ibáñez et al., 2006; Luck et al., 1996; Maki, et al., 1997; Vogel et al., 1998), such that an active task is not necessary for their elicitation, reducing task demands.

It should be noted that although we did not require a concurrent behavioral response, participants were told before the experiment that it was important for them to pay attention because questions about the pictures may be asked after the experiment. After the experiment each participant noted that some of the pictures matched the sound and when probed further, all participants were able to give word and environmental sound trial examples, indicating that they had indeed attended to the task.

2.5. EEG recording

EEG data was collected using a 64-electrode cap (Electro cap Inc.) according to the International 10–20 system. Tin electrodes were placed at the following locations (FP1, FPZ, FP2, AF3, AF4, F7, F5, F3, F1, FZ, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCZ, FC2, FC4, FC6, FT8, T7, C5, C3, C1, CZ, C2, C4, C6, T8, M1, TP7, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, M2, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, CB1, 01, OZ, O2, CB2). EOG was recorded from electrodes positioned on the outer canthi of each eye as well as above and below the left eye. 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 100 Hz. EEG gain was set to 20,000 and EOG gain set to 5000. Electrode impedances were kept below 5 kΩ.

2.6. EEG analysis

EEG was time locked to the auditory stimulus onset (spoken word or environmental sound) and epochs of 700 ms from auditory onset were averaged with a 100 ms pre-stimulus baseline. Trials containing eye movements, blinks, excessive muscle activity, or amplifier blocking were rejected trial-by-trial by off-line visual inspection before averaging (average rejection rate = 9%). Data for three subjects in the Word run, and one subject in the Environmental Sound run were removed due to consistent, pervasive broad-spectrum noise. To analyze potential differences in distributional effects across conditions while minimizing the number of total comparisons, we coded electrodes along two dimensions: Anteriority (anterior vs posterior) and Laterality (left, central, right), effectively dividing them into six regions: Left-Anterior (F7, F5, F3, FT7, FC5, FC3, T7, C5, C3), Left-Posterior (TP7, CP5, CP3, P7, P5, P3, PO7, PO5, CB1), Central-Anterior (F1, FZ, F2, FC1, FCZ, FC2, C1, CZ, C2), Central-Posterior (CP1, CPZ, CP2, P1, PZ, P2, PO3, POZ, PO4, O1, OZ, O2), Right-Anterior (F4, F6, F8, FC4, FC6, FT8, C4, C6, T8), and Right-Posterior (CP4, CP6, TP8, P4, P6, P8, PO6, PO8, CB2); to balance the number of electrodes in each region, the following electrodes were not included in the analysis: FP1, FPZ, FP2, AF3, AF4.

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 and Neville, 1990). Based on this prior work, and visual inspection of the grand average waveforms, we chose four time windows of interest: 200–300 ms, 300–400 ms, 400–500 ms, and 500–700 ms. For each sound type (Word and Environmental), mean amplitude voltage was computed separately for each condition (Match, Near Violations, and Far Violations) within the four time windows of interest. 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 fixed effects of Sound Type (Word or Environmental), Featural Similarity (Match, Near Violation, Far Violation), Anteriority (Anterior/Posterior), Laterality (Left/Center/Right), and their interactions. We report Type III *F*-tests for the main effects and interactions of these factors. The three levels of Featural Similarity (Match, Near Violations, Far Violations) were contrasted within each Sound Type (Words, Environmental Sounds) for significant Sound Type × Featural Similarity interactions. For these contrasts,

we report the regression coefficients (and standard error), *t*-values, *p*-values, and the 95% confidence interval. Note that the degrees of freedom are larger than in ANOVA approaches. 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

The overall ERP response for Words and Environmental Sounds was similar in morphology and scalp distribution (Figs. 2 and 3; Table 1). However, condition-specific differences were present. Broadly, the two sound types show a similar pattern of ERP components across the scalp starting with a central-anterior N100 peaking near 150 ms, followed by a large centrally distributed P200 at 275 ms and an anteriorly distributed N200 peaking around 350 ms. After the N200 the ERPs are largely characterized by slower and negative-going waves that last through the end of the recording epoch. Different patterns of condition-specific effects are found for Words and Environmental Sounds on these late negative-going waves. Words showed incremental amplitude difference between conditions throughout this period, whereas Environmental Sounds did not. Additionally, differences in the amplitude of Far Violations compared to Near Violations and Matches appear earlier for Environmental Sounds (~150 ms) than for Words (~300 ms). Here we present results within each of our four time windows of interest.

3.1. Time course analyses

3.1.1. 200–300 ms time window—This time window was characterized by a positive deflection in the waveform for all conditions. There were significant main effects of featural similarity [$F(2, 6747) = 132.39$, $p < 0.0001$], and sound condition [$F(1, 6751) = 298.85$, $p < 0.0001$], as well as a significant interaction [$F(2, 6747) = 35.64$, $p < 0.0001$]. Given the significant interaction, we examined contrasts between the levels of featural similarity separately for each sound type. For Words, Near Violations were significantly more positive than Matches, ($B = -0.59$ (0.13), $t(6747) = 4.45$, $p < 0.001$; 95% CI: [-0.85, -0.33]), and Far Violations ($B = -0.75$ (0.13), $t(6747) = 5.69$, $p < 0.0001$; 95% CI: [-1.01, -0.49]), and Matches and Far Violations were not significantly different ($B = -0.16$ (0.13), $t(6747) = 1.24$, $p = 0.21$; 95% CI: [-0.42, 0.10]). For Environmental Sounds, Near Violations were significantly more positive than Matches ($B = -0.91$ (0.13), $t(6747) = 7.21$, $p < 0.0001$; 95% CI: [-1.16, -0.67]) and Far Violations ($B = -2.22$ (0.13), $t(6747) = 17.61$, $p < 0.0001$; 95% CI: [-2.47, -1.97]). Matches were also significantly more positive than Far Violations ($B = -1.31$ (0.13), $t(6747) = 10.40$, $p < 0.0001$; 95% CI: [-1.56, -1.06]). There were no significant interactions between featural similarity and Laterality and Anteriority.

3.1.2. 300–400 ms time window—In this time window there were significant main effects of featural similarity [$F(2, 6747) = 205.46$, $p < 0.0001$], sound condition [$F(1, 6752) = 202.54$, $p < 0.001$], and a significant interaction between them [$F(2, 6747) = 39.16$, $p < 0.0001$]. Given the significant interaction, we examined contrasts between the levels of featural similarity separately for each sound type. For Words, a graded effect was observed

such that Far Violations were significantly more negative than both Matches ($B = -1.34$ (0.13), t (6747) = 8.66, $p < 0.0001$; 95% CI: [−1.39, −0.88]), and Near Violations ($B = -0.73$ (0.13), t(6747) = 5.53, $p < 0.0001$; 95% CI: [−0.98, −0.47]), and Near Violations were significantly more negative than Matches ($B = 0.41$ (0.13), t(6747) = 3.13, $p = 0.0018$; 95% CI: [0.15, 0.67]). Environmental Sounds elicited a different pattern such that Far Violations were significantly more negative than both Matches ($B = -2.18$ (0.12), t(6747) = 17.44, $p < 0.0001$; 95% CI: [−2.42, −1.93]), and Near Violations ($B = -2.30$ (0.12), t(6747) = 18.40, $p < 0.0001$; 95% CI: [−2.54, −2.05]), however Matches and Near Violations were not significantly different ($B = -0.12$ (0.12), t (6747) = 0.97, $p = 0.33$; 95% CI: [−0.37, 0.12]). There were no significant interactions between featural similarity and Laterality or Anteriority..

3.1.3. 400 – 500 ms time window—In this time window there were significant main effects of Featural Similarity [$F(2, 6747) = 351.24$, $p < 0.0001$], Sound Condition [$F(1, 6747) = 57.36$, $p < 0.0001$], and a significant interaction between them [$F(2, 6747) = 77.45$, $p < 0.0001$]. Similar to the 300– 400 ms time window, Words exhibited a graded effect such that Far Violations were marginally more negative than Near Violations ($B = -0.25$ (0.14), t(6747) = 1.79, $p = .07$; 95% CI: [−0.52, 0.024]) and significantly more negative than Matches ($B = -1.83$ (0.14), t (6747) = 13.15, $p < 0.0001$; 95% CI: [−2.10, −1.56]), and Near Violations were significantly more negative than Matches ($B = 1.58$ (0.14), t(6747) = 11.37, $p < 0.0001$; 95% CI: [1.31, 1.85]). In this time window, Environmental Sounds elicited a similar graded pattern such that Far Violations were significantly more negative than Near Violations ($B = -2.62$ (0.13), t(6747) = 19.8, $p < 0.0001$; 95% CI: [−2.88, −2.36]), and Matches ($B = -3.25$ (0.13), t(6747) = 24.5, $p < 0.0001$; 95% CI: [−3.51, −2.99]), and Near Violations were significantly more negative than Matches ($B = 0.62$ (0.13), t(6747) = 4.71, $p < 0.0001$; 95% CI: [0.36, 0.88]). There were no significant interactions between Featural Similarity and Laterality or Anteriority.

3.1.4. 500 – 700 ms time window—In this time window there were main effects of Featural Similarity [$F(2, 6747) = 269.81$, $p < .0001$], Sound Condition [$F(1, 6751) = 29.55$, $p < 0.0001$], and a significant interaction between them [$F(2, 6747) = 58.38$, $p < 0.0001$]. Consistent with the previous two time windows, Words demonstrated a graded amplitude pattern based on Featural Similarity such that Far Violations were significantly more negative than Near Violations ($B = -0.55$ (0.15), t(6747) = 3.59, $p = 0.0003$; 95% CI: [−0.86, −0.25]), and Matches ($B = -2.29$ (0.15), t(6747) = 14.77, $p < 0.0001$; 95% CI: [−2.59, −1.98]), and Near Violations were significantly more negative than Matches ($B = 1.73$ (0.15) t(6747) = 11.19, $p < .0001$; 95% CI: [1.43, 2.03]). Environmental Sounds did not exhibit a graded effect: Far Violations were significantly more negative than Matches ($B = -2.59$ (0.15), t(6747) = 17.57, $p < .0001$; 95% CI: [−2.88, −2.30]) and Near Violations ($B = -2.69$ (0.15), t(6747) = 18.24, $p < .0001$; 95% CI: [−2.98, −2.40]), however the Matches and Near Violations were not significantly different ($B = -0.10$ (0.15), t (6747) = 0.67, $p = 0.50$; 95% CI: [−0.39, 0.19]). There were no significant interactions between Featural Similarity and Laterality or Anteriority.

4. Discussion and conclusion

The primary aim of this study was to determine whether words and environmental sounds are organized similarly in semantic memory, and whether environmental sounds appear to be organized in memory according to featural similarity. We systematically varied the degree of featural similarity between an auditory stimulus and a preceding pictorial context for both sound classes. Consistent with previous work using similar methods, the overall ERP response for both sound types was similar in morphology and scalp distribution (Figs. 2 and 3). However, condition-specific differences were present. Far Violations (within-category but relatively unrelated to the preceding context), exhibited significant N400 effects for words and environmental sounds throughout all time windows of interest. This replicates and extends previous work showing that when put into similar context-dependent situations, environmental sounds elicit N400 peaks to between category violations similar to those elicited by auditory or visual words (Cummings et al., 2008; van Petten and Rheinfelder, 1995; Plante et al., 2000, Van Petten and Senkfor, 2000).

The finding that N400 effects to words and environmental sounds have similar scalp topographies is consistent with previous work using a similar cross-modal ERP priming paradigm (Cummings et al., 2006, 2008; Cummings and Èeponiene, 2010; Orgs et al., 2007). Other previous studies that report differences in scalp topography between words and environmental sounds used uni-modal auditory stimulus presentation (spoken words primed environment sounds and environmental sounds primed spoken words; Van Petten and Rheinfelder, 1995). Once differences in presentation modality are taken into account, the current scalp topography results are indeed consistent with the extant literature.

Specific differences were also observed for each distinct sound type. Most relevant to the primary goal of the current paper are the differences in N400 amplitude for each sound type. For words, consistent with our prediction, we replicated previous research demonstrating gradedness in N400 amplitude modulated by featural similarity: both Near and Far Violations exhibited significant N400 effects, with Far Violations exhibiting greater N400 amplitudes than Near Violations. This graded pattern was consistent, starting in the early N400 time window (300–400 ms) and continuing throughout the epoch. In contrast to words, environmental sounds showed a pattern in which the Far Violations but not the Near Violations elicited significant N400 effects for both the early N400 time window (300–400 ms) and the late N400 time window (500–700 ms). The pattern of effects exhibited by environmental sounds is consistent with the alternative prediction based on behavioral work demonstrating that recognition (accuracy and response time) for environmental sound-object pairs is more affected by changes in feature likeness than is the recognition of word-object pairs (Saygin et al., 2005). Thus our findings lend supp ort to the idea that the processing of environmental sounds is less sensitive to subtle changes in semantic features, and is organized in terms of a more general feature match between sounds and referents.

Environmental sounds did exhibit graded N400 amplitudes between 400–500 ms, although this effect was clearly shorter in duration than the graded effect observed for words. One possible interpretation for the relatively short-lived gradedness for environmental sounds from 400–500 ms is that the effect is an artifact of overlapping components. A slight

positive deflection for the environmental sound Match condition in this time window would be consistent with a P300 component (Polich, 2012). Although this experiment did not require an explicit, task relevant response, and therefore not constructed to elicit P300 effects, the a priori probability of getting a Match was lower than the probability of getting a Violation (combined Near and Far conditions). Consistent with this interpretation, there is precedence in the literature for P300 effects being elicited in a passive listening paradigm (Bennington amd Polich, 1999). Thus, the fleeting graded effect between 400–500 ms for environmental sounds may be in part due to the Match condition demonstrating a greater positive deflection, which briefly distinguished the amplitudes of Near Violations and Matches. Indeed we observe such a putative positive deflection for words as well, however because Matches were more positive than both Violation conditions throughout the negative- going epochs, the positive deflection to Matches did not change the existing pattern of effects.

We also found differences in how feature similarity affected the processing of words and environmental sounds in the 200–300 ms time window: Matches were significantly more positive than Far Violations for environmental sounds but not words. We find two possible explanations for this early effect of featural similarity. Firstly, it may be due to low-level acoustic differences between the stimuli in the different conditions. However, we find this interpretation unlikely because the sounds used in the three conditions were highly overlapping, with virtually all stimuli present in all three conditions (e.g., *Ringing* is an environmental sound present in all three conditions, as is the corresponding word "telephone"; see Appendix). Another possible explanation is that these results reflect an early onset of the N400 component to the Far Violations for environmental sounds, causing them to be more negative in this early time window. As discussed above (see Section 1), the effects of semantic priming start earlier for auditory stimuli (Holcomb and Neville, 1990), and therefore may have affected the relative amplitudes of the featural similarity conditions in the preceding positive-going deflection. Moreover, previous research has shown that N400 priming effects appear earlier for environmental sounds than for words (Cummings et al., 2006; Orgs et al., 2006), consistent with the apparent lack of an early-onset N400 for words in this time window in our results.

Overall, our results suggest a fine-grained graded organizational structure for words based on featural similarity, with only a coarse distinction apparent for the relation between environmental sounds and concepts. However, there are several potential alternative hypotheses for the currents results: acoustic similarity between environmental sounds, precise encoding of lexical items, and differences in listener's familiarity with the sounds. Here we will discuss each alternative hypothesis in detail.

Although environmental sounds obtain their meaning through a causal production and are therefore, in principle bound to their sound source, in practice it may be quite difficult to determine the associated concept from hearing the sound alone (Ballas et al., 1986). It has been shown that there are more similar sounding environmental sounds that have different sources than there are similar sounding words with different meanings. For example, the environmental sound, "click–click" is ambiguous, and can be produced by a pen, a light switch, or many other possible referents (Ballas, 1993). Indeed people show more variation

in their ability to recognize and identify environmental sounds than words (Gygi, 2001). Therefore it is possible that acoustic similarities between environmental sounds may be the interfering factor. That is, environmental sounds may be more coarsely organized in semantic memory than words because acoustic similarities between different environmental sounds causes an initial misread in the raw acoustic signal and results in a semantic misinterpretation. Relatedly, the acoustic stimuli for the environmental sound condition was numerically longer than the word condition (however, this difference did not reach significance). Therefore it is possible that the generation of the N400 responses was affected by the average duration of the two types of stimuli.

Another alternative interpretation is centered on lexical items being more precisely encoded. Unlike environmental sounds whose raw acoustic signal is likely directly linked to the semantic representation, word recognition undergoes multiple consecutive processing stages that translate a raw acoustic signal to a meaningful symbolic unit. Spoken words are first processed as a nonlinguistic auditory signal, then progress through a series of linguistic specific processing stages (phonetic and phonemic), and finally, higher-level brain structures compute semantic relevance (Frauenfelder and Tyler 1987; Indefrey and Levelt, 2004). The extra lexical encoding stages inherent in processing words may seek to further distinguish the raw acoustic input as referring to one possible referent. This precise level of encoding that is allowed in a relational mental lexicon (i.e., semantic and phonological relations), may not be possible for environmental sounds.

Finally, another interpretation highlights the possible differences in *a priori* levels of familiarity and exposure people have with words vs. environmental sounds. Although we pretested the materials in a group of adults to try to ensure that the environmental sounds were highly familiar, it is difficult to say with certainty whether there are differences in familiarity (and exposure) between these words and environmental sounds that could be driving the different ERP patterns. That is, participants could have more *a priori* exposure to the word stimuli than the environmental sound stimuli, and therefore, have had more time to semantically organize words in long-term memory. For the current study we included highly familiar words and environmental sounds to (a) increase the probability that participants have had long-term exposure to the auditory stimuli and, (b) ensure replication of previous findings of gradedness in N400 amplitudes for familiar words (Federmeier et al., 2002; Federmeier and Kutas, 1999, 2002; Ibáñez et al., 2006; Torkilson et al., 2006).

To our knowledge, the evidence presented here is the first to show differential semantic memory organization between linguistic and meaningful nonlinguistic auditory information. An important question for future research is whether behavioral measures would also be able to reveal the observed difference (e.g. graded vs. non-graded) in memory structure for the two types of auditory stimuli. The present study furthers our understanding of how the brain organizes meaningful auditory information and the relation between language and cognition. Additionally, this work has implications for broader theoretical accounts regarding whether the ability to procure meaning from words is subserved by semantic resources that are specific to language, or more global, cross-domain memory stores.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Amunts K, Zilles K. Architecture and organizational principles of Broca's area. Trends. Cogn. Sci. 2012; 16(8):418–426. [PubMed: 22763211]

Aramaki M, Marie C, Kronland-Martinet R, Ystad S, Besson M. Sound categorization and conceptual priming for nonlinguistic and linguistic sounds. J. Cogn. Neurosci. 2010; 22(11):2555–2569. [PubMed: 19929328]

Ballas J. Common factors in the identification of an assortment of brief everyday sounds. J. Exp. Psychol.: Hum. Percep. Perform. 1993; 19(2):250–267.

Ballas J, Howard J. Interpreting the language of environmental sounds. Environ. Behav. 1987; 19(1): 91–114.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. Cerebral Cortex. 2000; 10(5):512–528. [PubMed: 10847601]

Gygi, B. Factors in the identification of environmental sounds (Unpublished doctoral dissertation). Bloomington: Indiana University; 2001.

Chen YC, Spence C. Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. J. Exp. Psychol. Hum. Percep. Perform. 2011; 37(5):1554.

Cummings A, Èeponiene R. Verbal and nonverbal semantic processing in children with developmental language impairment. Neuropsychologia. 2010; 48(1):77–85. [PubMed: 19698728]

Cummings A, Èeponiene R, Dick F, Saygin AP, Townsend J. A developmental ERP study of verbal and non-verbal semantic processing. Brain Res. 2008; 1208:137. [PubMed: 18387601]

Cummings A, Èeponiene R, Koyama A, Saygin AP, Townsend J, Dick F. Auditory semantic networks for words and natural sounds. Brain Res. 2006; 1115(1):92–107. [PubMed: 16962567]

Cycowicz, Yael M.; Friedman; David. Effect of sound familiarity on the event-related potentials elicited by novel environmental sounds. Brain Cogn. 1998; 36(1):30–51. [PubMed: 9500881]

Daltrozzo J, Schön D. Conceptual processing in music as revealed by N400 effects on words and musical targets. J. Cogn. Neurosci. 2009; 21(10):1882–1892. [PubMed: 18823240]

Federmeier KD, Kutas M. A rose by any other name: Long-term memory structure and sentence processing. J. Mem. Lang. 1999; 41(4):469–495.

Federmeier KD, McLennan DB, Ochoa E, Kutas M. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. Psychophysiology. 2002; 39(2):133–146. [PubMed: 12212662]

Frauenfelder UH, Tyler LK. The process of spoken word recognition: an introduction. Cognition. 1987; 25(1):1–20. [PubMed: 3581722]

Frey A, Aramaki M, Besson M. Conceptual priming for realistic auditory scenes and for auditory words. Brain Cogn. 2014; 84(1):141–152. [PubMed: 24378910]

Griffith TD, Flees A, Green GG. Disorders of human complex sound processing. Neurocase. 1999; 5(5):365–378.

Griffiths TD, Warren JD, Scott SK, Nelken I, King AJ. Cortical processing of complex sound: a way forward? TRENDS in Neurosciences. 2004; 27(4):181–185. [PubMed: 15046876]

Hocking J, Dzafic I, Kazovsky M, Copland DA. NESSTI: norms for environmental sound stimuli. PloS One. 2013; 8(9):73382.

Holcomb PJ, Neville HJ. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Lang. Cogn. Process. 1990; 5(4):281–312.

Hubel DH, Wiesel TN. Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. J. Comparat. Neurol. 1972; 146(4):421–450.

Humphries C, Willard K, Buchsbaum B, Hickok G. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. Neuroreport. 2001; 12(8):1749–1752. [PubMed: 11409752]

Ibáñez A, López V, Cornejo C. ERPs and contextual semantic discrimination: degrees of congruence in wakefulness and sleep. Brain Lang. 2006; 98(3):264–275. [PubMed: 16782185]

Indefrey P, Levelt WJ. The spatial and temporal signatures of word production components. Cognition. 2004; 92(1):101–144. [PubMed: 15037128]

Kay P. Taxonomy and semantic contrast. Language. 1971:866–887.

Kimura D. From ear to brain. Brain Cogn. 2011; 76(2):214–217. [PubMed: 21236541]

Knox C, Kimura D. Cerebral processing of nonverbal sounds in boys and girls. Neuropsychologia. 1970; 8(2):227–237. [PubMed: 5535486]

Koelsch S. Neural substrates of processing syntax and semantics in music. Curr. Opin. Neurobiol. 2005:207–212. [PubMed: 15831404]

Koelsch S, Kasper E, Sammler D, Schulze K, Gunter T, Friederici AD. Music, language and meaning: Brain signatures of semantic processing. Nat. Neurosci. 2004; 7(3):302–307. [PubMed: 14983184]

Kutas M, Federmeier KD. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 2011; 62:621–647. [PubMed: 20809790]

Kutas M, Hillyard SA. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science. 1980; 207(4427):203–205. [PubMed: 7350657]

Kutas M, Hillyard SA. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. 1983; 11(5):539–550.

Kutas M, Van Petten C. Event-related brain potential studies of language. Adv. Psychophysiol. 1988; 3:139–187.

Kutas M, Van Petten C. Psycholinguistics electrified. Handbook of Psycholinguistics. 1994:83–143.

Leech R, Saygin AP. Distributed processing and cortical specialization for speech and environmental sounds in human temporal cortex. Brain Lang. 2011; 116(2):83–90. [PubMed: 21167584]

Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. Distinct cortical pathways for processing tool versus animal sounds. The Journal of Neuroscience. 2005; 25(21):5148–5158. [PubMed: 15917455]

Luck SJ, Vogel EK, Shapiro KL. Word meanings can be accessed but not reported during the attentional blink. Nature. 1996; 383(6601):616–618. [PubMed: 8857535]

Maki WS, Frigen K, Paulson K. Associative priming by targets and distractors during rapid serial visual presentation: does word meaning survive the attentional blink? J. Exp. Psychol.: Hum. Percep. Perform. 1997; 23(4):1014.

McRae K, Cree GS, Seidenberg MS, McNorgan C. Semantic feature production norms for a large set of living and nonliving things. Behav. Res. Methods. 2005; 37(4):547–559. [PubMed: 16629288]

Murphy GL, Hampton JA, Milovanovic GS. Semantic memory redux: an experimental test of hierarchical category representation. J. Mem. Lang. 2012; 67(4):521–539. [PubMed: 23243336]

Newman AJ, Tremblay A, Nichols ES, Neville HJ, Ullman MT. The influence of language proficiency on lexical semantic processing in native and late learners of English. J. Cogn. Neurosci. 2012; 24(5):1205–1223. [PubMed: 21981676]

Noppeney U, Josephs O, Hocking J, Price CJ, Friston KJ. The effect of prior visual information on recognition of speech and sounds. Cereb. Cortex. 2008; 18(3):598–609. [PubMed: 17617658]

Orgs G, Lange K, Dombrowski JH, Heil M. Conceptual priming for environmental sounds and words: an ERP study. Brain Cogn. 2006; 62(3):267–272. [PubMed: 16793186]

Orgs G, Lange K, Dombrowski JH, Heil M. Is conceptual priming for environmental sounds obligatory? Int. J. Psychophysiol. 2007; 65(2):162–166. [PubMed: 17448556]

Orgs G, Lange K, Dombrowski JH, Heil M. N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. Neurosci. Lett. 2008; 436(2):133–137. [PubMed: 18378076]

Özcan E, Egmond RV. The effect of visual context on the identification of ambiguous environmental sounds. Acta Psychol. 2009; 131(2):110–119.

Paczynski M, Kuperberg GR. Multiple influences of semantic memory on sentence processing: distinct effects of semantic relatedness on violations of real-world event/state knowledge and animacy selection restrictions. J. Mem. Lang. 2012; 67(4):426–448. [PubMed: 23284226]

Perrin F, Garcýa-Larrea L. Modulation of the N400 potential during auditory phonological/semantic interaction. Cogn. Brain Res. 2003; 17(1):36–47.

Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, Di Russo F. Separate neural systems for processing action-or non-action-related sounds. Neuroimage. 2005; 24(3):852–861. [PubMed: 15652320]

Plante E, Petten CV, Senkfor AJ. Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. Neuropsychologia. 2000; 38(13):1669–1684. [PubMed: 11099725]

Polich J. Neuropsychology of P300. Oxford Handbook of Event-related Potential Components. 2012:159–188.

Polich J. Semantic categorization and event-related potentials. Brain Lang. 1985; 2:304–321. [PubMed: 4084767]

Bennington JY, Polich J. Comparison of P300 from passive and active tasks for auditory and visual stimuli. Int. J. Psychophysiol. 1999; 34(2):171–177. [PubMed: 10576401]

Price C, Thierry G, Griffiths T. Speech-specific auditory processing: where is it? Trends Cogn. Sci. 2005; 9(6):271–276. [PubMed: 15925805]

Renvall H, Formisano E, Parviainen T, Bonte M, Vihla M, Salmelin R. Parametric merging of MEG and fMRI reveals spatiotemporal differences in cortical processing of spoken words and environmental sounds in background noise. Cereb. Cortex. 2012; 22(1):132–143. [PubMed: 21613467]

Richards JE. Attention affects the recognition of briefly presented visual stimuli in infants: an ERP study. Dev. Sci. 2003; 6(3):312–328. [PubMed: 16718304]

Rosch E, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P. Basic objects in natural categories. Cogn. Psychol. 1976; 8(3):382–439.

Sajin SM, Connine CM. Semantic richness: The role of semantic features in processing spoken words. J. Mem. Lang. 2014; 70:13–35.

Saygin AP, Dick F, Bates E. An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. Behav. Res. Methods. 2005; 37(1):99–110. [PubMed: 16097349]

Schirmer A, Soh YH, Penney TB, Wyse L. Perceptual and conceptual priming of environmental sounds. J. Cogn. Neurosci.e. 2011; 23(11):3241–3253.

Schneider TR, Engel AK, Debener S. Multisensory identification of natural objects in a two-way cross-modal priming paradigm. Exp. Psychol. 2008; 55(2):121–132. [PubMed: 18444522]

Schön D, Ystad S, Kronland-Martinet R, Besson M. The evocative power of sounds: Conceptual priming between words and nonverbal sounds. J. Cogn. Neurosci. 2010; 22(5):1026–1035. [PubMed: 19583472]

Shafiro V, Gygi B. How to select stimuli for environmental sound research and where to find them. Behav. Res. Methods, Instrum. Comput. 2004; 36(4):590–598. [PubMed: 15641404]

Steinbeis N, Koelsch S. Comparing the processing of music and language meaning using EEG and FMRI provides evidence for similar and distinct neural representations. PLoS One. 2008; 3(5):2226.

Stuart GP, Jones DM. Priming the identification of environmental sounds Q. J. Exp. Psychol. 1995; 48(3):741–761.

Tanaka K. Representation of visual features of objects in the inferotemporal cortex. Neural Netw. 1996; 9(8):1459–1475. [PubMed: 12662545]

Thierry G, Giraud AL, Price C. Hemispheric dissociation in access to the human semantic system. Neuron. 2003; 38(3):499–506. [PubMed: 12741995]

Tranel D, Damasio H, Eichhorn GR, Grabowski T, Ponto LL, Hichwa RD. Neural correlates of naming animals from their characteristic sounds. Neuropsychologia. 2003; 41(7):847–854. [PubMed: 12631534]

Van Petten C, Rheinfelder H. Conceptual relationships between spoken words and ES: event-related brain potential measures. Neuropsychologia. 1995; 33(4):485–508. [PubMed: 7617157]

Vogel EK, Luck SJ, Shapiro KL. Electrophysiological evidence for a post-perceptual locus of suppression during the attentional blink. J. Exp. Psychol.: Hum. Percept. Perform. 1998; 24(6): 1656. [PubMed: 9861716]
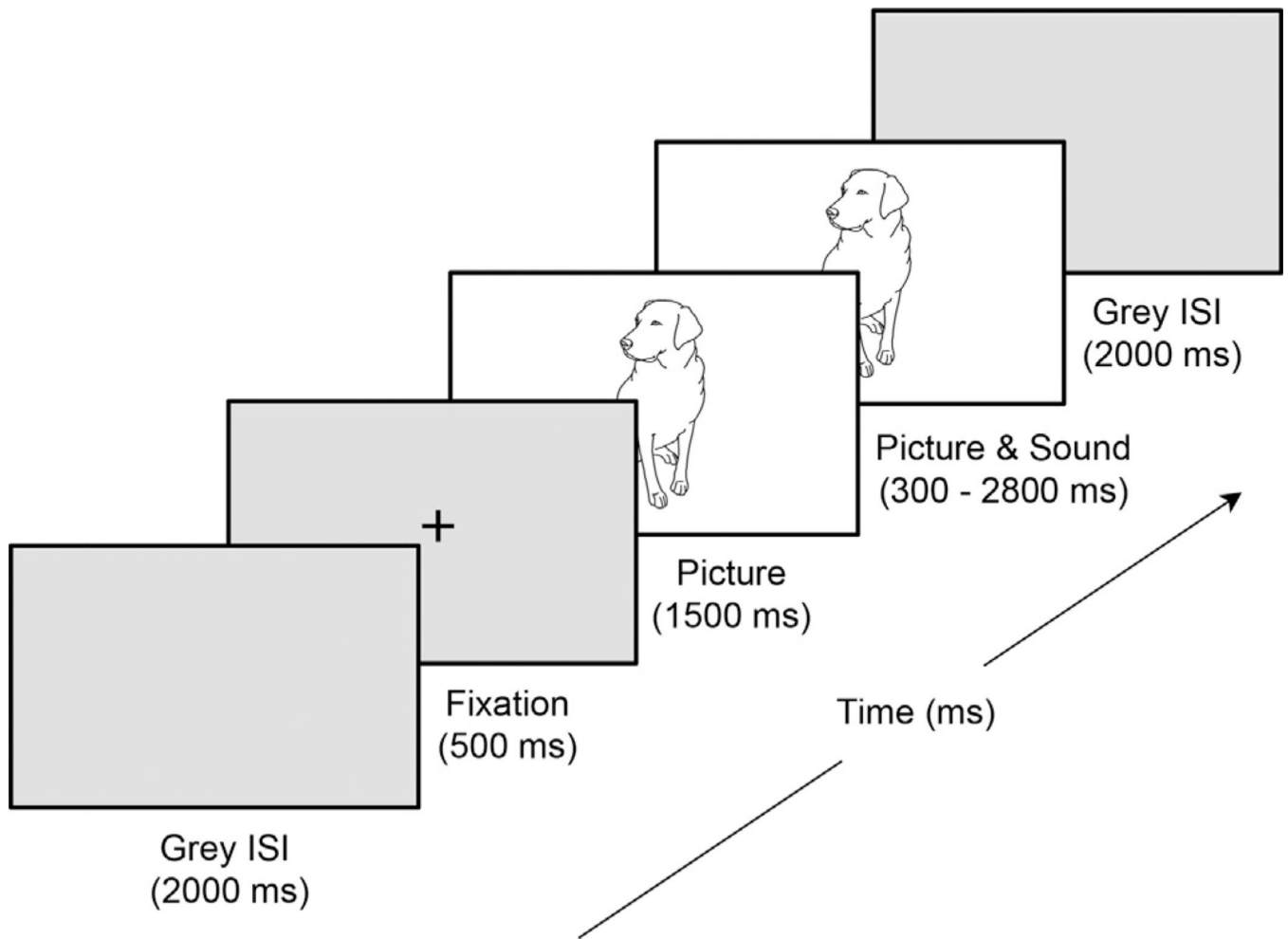
Fig. 1.
Schematic of a single trial. For each sound type (Word and Environmental Sound), participants were presented with a line drawing of a familiar concept for 1500 ms before hearing a sound (300–2800 ms duration) from one of three conditions (Matches, Near Violation, Far Violation). The picture disappeared at the offset of the sound.
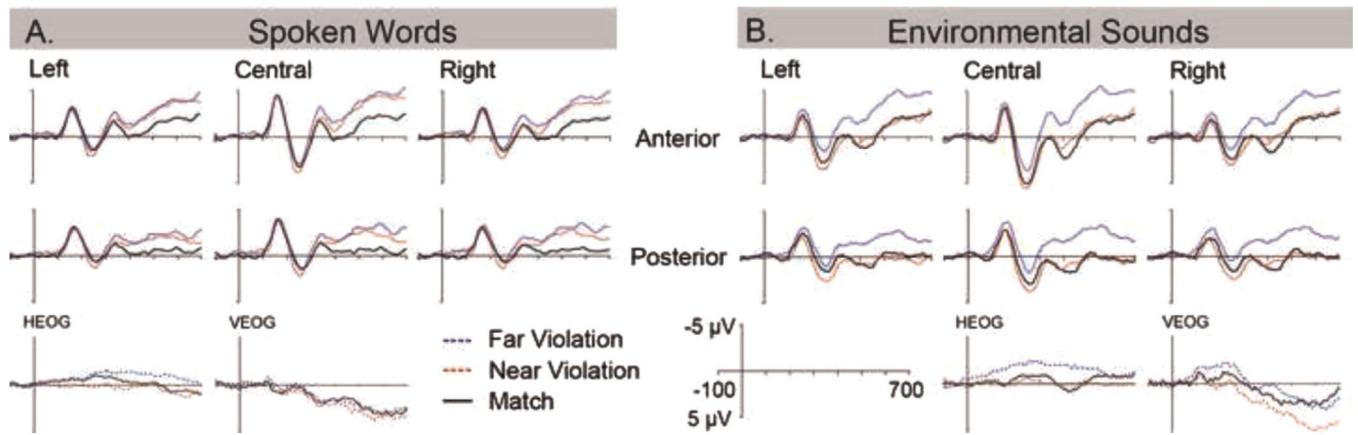
Fig. 2.
Grand average ERP waveforms for the three levels of featural similarity for six regions –
left-anterior, left-posterior, central-anterior, central-posterior, right-anterior, right-posterior –
for Words (A) and Environment Sounds (B). ERPs from vertical electro-oculogram (VEOG)
and horizontal electro-oculogram (HEOG) channels are included for illustrative purposes.

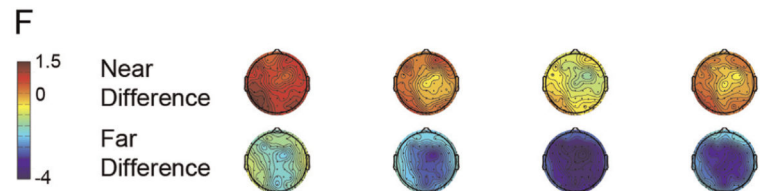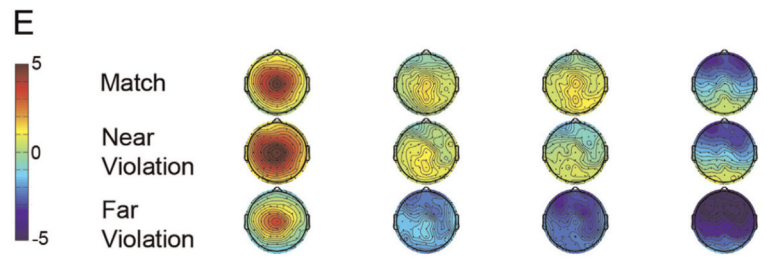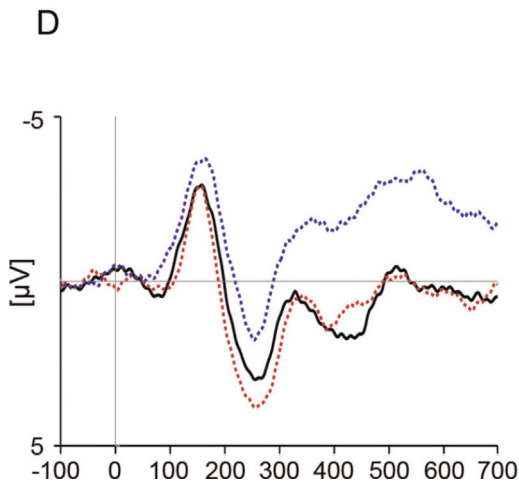## Spoken Words
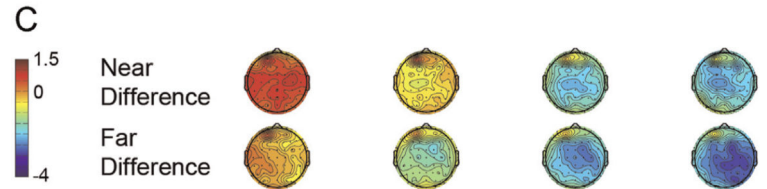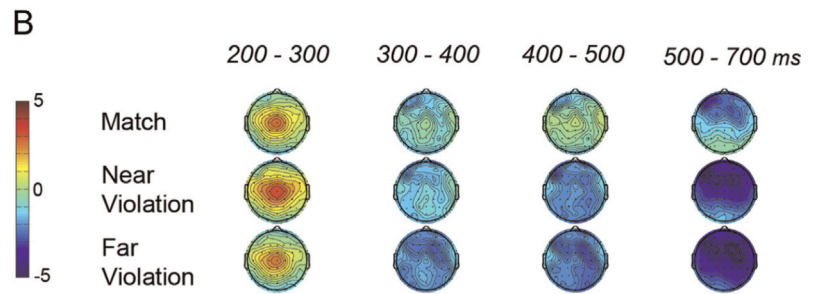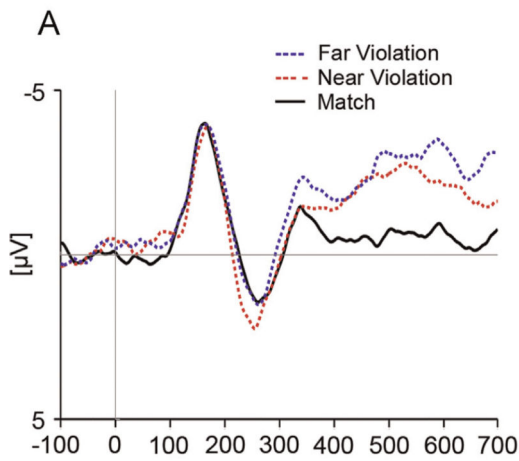
Environmental Sounds

Fig. 3.
ERP waveforms and voltage maps. Effect of featural similarity shown at the central-posterior region for Words (A) and Environmental Sounds (C). Voltage maps show average voltage for Words (B) and Environmental Sounds (E), and average voltage difference (measured as Violations – Matches) for Words (C) and Environmental Sounds (F) between two fixed latencies.

Table 1

Mean Voltage (μV) for each Featural Similarity Condition for each Sound type, across all electrodes, at four time windows of interest. Standard error is shown in parentheses.

Time windows (ms)	Conditions							
	Words				**Environmental Sounds**			
	Match	**Near**	**Far**		**Match**	**Near**	**Far**	
200–300	0.65 (0.10)	1.24 (0.08)	0.49 (0.09)		1.98 (1.3)	2.89 (1.2)	2.89 (1.2)	
300–400	−0.75 (0.10)	−1.16 (0.07)	− 1.88 (0.09)		− 0.53 (0.15)	− 0.65 (0.11)	− 0.65 (0.11)	
400–500	−29 (0.10)	−1.87 (0.08)	−2.12 (0.10)		−0.78 (0.17)	−0.16 (0.12)	−2.46 (0.17)	
500–700	−1.26 (0.10)	−2.99 (0.09)	−3.54 (0.11)		−1.08 (0.18)	−0.98 (0.14)	−0.98 (0.14)	