Sound Categorization and Conceptual Priming for Nonlinguistic and Linguistic Sounds

Mitsuko Aramaki[1,2], Céline Marie[1,2], Richard Kronland-Martinet[3], Sølvi Ystad[3], and Mireille Besson[1,2]

Abstract

■ The aim of these experiments was to compare conceptual priming for linguistic and for a homogeneous class of nonlinguistic sounds, impact sounds, by using both behavioral (percentage errors and RTs) and electrophysiological measures (ERPs). Experiment 1 aimed at studying the neural basis of impact sound categorization by creating typical and ambiguous sounds from different material categories (wood, metal, and glass). Ambiguous sounds were associated with slower RTs and larger N280, smaller P350/P550 components, and larger negative slow wave than typical impact sounds. Thus, ambiguous sounds were more difficult to categorize than typical sounds. A category membership task was used in Experiment 2. Typical sounds were followed by sounds from the same or from a different category or by ambiguous sounds. Words were followed by words, pseudowords, or nonwords. Error rate was highest for ambiguous sounds and for pseudowords and both elicited larger N400-like components than same typical sounds and words. Moreover, both different typical sounds and nonwords elicited P300 components. These results are discussed in terms of similar conceptual priming effects for nonlinguistic and linguistic stimuli. ■

INTRODUCTION

Word processing is specific in that the sequences of phonemes that form spoken words (or the sequence of graphemes that form written words) are not meaningful by themselves but acquire meaning through the process of double articulation. Thus, there is generally no direct relationship between the sound (or form) of a word and its meaning (de Saussure, 1916). By contrast, a causal relationship exists between the perceptual characteristics of environmental sounds (e.g., broken glass) and the meaning that is derived from them (e.g., a glass was broken; Ballas, 1993; Ballas & Howard, 1987). On the basis of such differences, one may expect words and environmental sounds to be processed differently. However, results of several experiments that have used the ERP method to address this question argue in favor of the similarities rather than the differences in the processing of words and environmental sounds (Orgs, Lange, Dombrowski, & Heil, 2006, 2007, 2008; Cummings et al., 2006; Plante, Van Petten, & Senkfor, 2000; Van Petten & Rheinfelder, 1995).

One of the first studies of conceptual priming was conducted by Van Petten and Rheinfelder (1995). In their first experiment, environmental sounds were used as primes and related words, unrelated words, and pseudowords were used as targets. Participants were asked to decide whether the target stimulus was a word or not (lexical decision task). Results showed higher error rate for pseudowords than for words and faster RTs for related than for unrelated words. Thus, these results demonstrated conceptual priming between environmental sounds and words. To examine the neural basis of this effect, they took advantage of the N400 component (Kutas & Hillyard, 1980) to compare priming when the prime is an environmental sound and the target a related or an unrelated printed word and vice versa. Results revealed that conceptual priming, as reflected by the N400 effect (i.e., the difference between unrelated and related targets), was very similar for environmental sounds and for words. This result argues in favor of the similarity of the neural processes involved in computing the meaning of words and environmental sounds. However, although the typical, slightly larger over the right than left hemisphere distribution of the N400 effect was found for target words (Kutas, Van Petten, & Besson, 1988; Kutas & Hillyard, 1982), the N400 effect to environmental sounds was larger over the left hemisphere. Similar interhemispheric differences in the N400 effect were reported in a subsequent study (Plante et al., 2000) in which priming for pairs of line drawing and environmental sounds, on one side, and for pairs of printed and spoken words, on the other side, was compared using the same task as in Van Petten and Rheinfelder (1995).

More recently, Orgs et al. (2006) used printed words as primes followed by environmental sounds as targets (or vice versa). Primes and targets were semantically related or

[1]CNRS—Institut de Neurosciences Cognitives de la Méditerranée, Marseille Cedex, France, [2]Université Aix-Marseille, Marseille Cedex, France, [3]CNRS—Laboratoire de Mécanique et d'Acoustique, Marseille Cedex, France

unrelated, and participants were asked to decide whether the words and environmental sounds fitted together or not. Results showed slower RTs and larger N400 amplitude for unrelated than for related targets. Moreover, the N400 effect showed an earlier onset for environmental sounds than for words over parieto-occipital sites. In subsequent studies with similar design and stimuli, participants were asked to perform either a physical task or a semantic task (Orgs et al., 2007, 2008). In both tasks, the authors found an N400 effect that was taken to reflect the automatic processes that mediate analysis of sound meaning. Finally, Cummings et al. (2006) also used a cross-modal priming design with pictures presented as primes and related or unrelated spoken words, environmental sounds, or nonmeaningful sounds (defined by the authors as "computer-generated sounds that were not easily associated with any concrete semantic concept"; Cummings et al., 2006, p. 104) presented as targets. Participants were asked to decide whether the two stimuli matched or mismatched. They reported an N400 effect for both words and environmental sounds but not for nonmeaningful sounds. Again, the N400 effect had an earlier onset for environmental sounds than for words, but in contrast with Orgs et al. (2006), the N400 effect was larger for environmental sounds than for words over frontal sites (F3/F4). Moreover, in contrast to Plante et al. (2000) and Van Petten and Rheinfelder (1995), they reported no interhemisphere differences. Findings from an experiment using event-related desynchronization to compare words and environmental sounds suggest the involvement of left-lateralized phonological and semantic processes for words and of distributed processes in both hemispheres for environmental sounds (Lebrun et al., 2001). On the basis of these results, Lebrun et al. (2001) suggested a common semantic system for both words and environmental sounds but with more specific perceptual processing for the later.

Other studies have also examined conceptual priming between music and language. Results have shown the occurrence of an N400 component to unrelated visual word targets when primes were long musical excerpts (several seconds; Koelsch et al., 2004). Recently, Daltrozzo and Schön (2009) used either words or short musical excerpts (1 sec) as primes and targets and found an N400 effect in both cases. However, the scalp distribution of the N400 effect was temporal for musical excerpts and parietal for visual word targets. To our knowledge, only one study has examined priming effects for pairs of musical excerpts presented both as prime and as target stimuli (Frey et al., 2009). Although the musical excerpts used as primes conveyed specific concepts, the musical excerpts used as targets either conveyed the same concept as the prime (congruous) or started with the same concept but shifted midstream into another concept (incongruous). Results showed an N400-like component, which was largest over right frontal regions, to incongruous musical excerpts in nonmusicians.

Taken together, results of these studies, which used different experimental designs, stimuli, and tasks (i.e., lexical decision task, matching tasks, physical or semantic priming), nevertheless concur in showing that environmental sounds or musical excerpt targets that are unrelated to word or picture primes elicit increased negative components in the N400 latency band compared with related targets. However, the scalp distribution of these effects varies either between hemispheres (Lebrun et al., 2001; Plante et al., 2000; Van Petten & Rheinfelder, 1995) or along the anterior-posterior dimension (Daltrozzo & Schön, 2009; Cummings et al., 2006; Orgs et al., 2006). Thus, whether conceptual priming for environmental sounds, music, and words rely on similar or different processes still remains an open issue.

The Present Studies

In the different studies summarized above (except for Frey et al., 2009), some form of cross-modal priming was always used between pictures, printed or spoken words, musical excerpts, on one side, and nonlinguistic sounds (environmental sounds, musical excerpts), on the other side. As a consequence, the presence of words or pictures in the experimental design may have encouraged the use of linguistic encoding strategies. Thus, the possibility remains that, within such experimental contexts, participants associated a verbal label to each sound thereby explaining the similarity of the N400 for words and environmental sounds. In a recent study, Schön, Ystad, Kronland-Martinet, and Besson (2009) also used a cross-modal priming design but they tried to minimize linguistic mediation by presenting sounds with no easily identifiable sound sources (i.e., a verbal label could not easily be associated to the sounds). Again, larger N400 components were found for targets (sounds or words) that were unrelated to the primes (words or sounds). However, this cross-modal design still involved words as prime or target stimuli. Therefore, our first aim was to reduce the influence of linguistic mediation by using only nonlinguistic sounds as prime and target. Moreover, in previous studies, the set of nonlinguistic sounds that were used were often very diverse and nonhomogeneous (e.g., animal or human nonspeech sounds, instrumental sounds, and everyday life sounds). As these different types of sounds may engage different processes, our second aim was to use only one homogeneous class of environmental sounds: impact sounds from wood, metal, or glass. Finally, to our knowledge, priming effects including only linguistic primes and linguistic targets on one side and nonlinguistic primes and nonlinguistic targets on the other side have never been directly compared within subjects. Thus, our third aim was to directly compare within-subjects conceptual priming for impact sounds and for linguistic sounds by using the same task with both types of stimuli. However, to use a priming design including only nonlinguistic sounds as stimuli, we first needed to create typical sounds from different impact sound categories (i.e., wood, metal, glass) and ambiguous sounds. The aims

of Experiment 1 were to create such sounds and to study the neural basis of impact sound categorization.

EXPERIMENT 1: SOUND CATEGORIZATION

To create typical and ambiguous impact sounds, we used a morphing technique. First, sounds from three material categories (wood, metal, and glass) were recorded, analyzed, and resynthesized using an analysis–synthesis method (Aramaki & Kronland-Martinet, 2006) to generate realistic synthetic sounds. Second, sound continua were created that simulate progressive transitions between sounds from different materials (i.e., wood–metal, wood–glass, and metal–glass continua). Although sounds at extreme positions on the continua were synthesized to be as similar as possible to natural sounds, sounds at intermediate positions were synthesized by interpolating the acoustic parameters characterizing sounds at extreme positions. They were consequently ambiguous (e.g., neither wood nor metal). Sounds from the different continua were randomly presented, and participants were asked to categorize each sound as wood, metal, or glass. If sounds at extreme positions of the continua are indeed perceived as typical exemplars of their respective categories, they should be categorized faster and with lower error rate than sounds at intermediate positions on the continua.

Little is known about the neural basis of impact sound perception. To investigate this issue, we also recorded ERPs while participants performed the categorization task. Results of studies on the categorization of nonspeech stimuli have shown that the amplitude of the N200–P300 complex, which typically follows the N100–P200 exogenous complex, is influenced by the difficulty of the categorization task: The N200 component is larger and the P300 component is smaller to stimuli that are more difficult to categorize (Donchin & Coles, 1988; Ritter, Simson, & Vaughan, 1983; Donchin, 1981; Ritter, Simson, Vaughan, & Friedman, 1979; Simson, Vaughan, & Ritter, 1977). Thus, if ambiguous sounds are more difficult to categorize than typical sounds because they are composed of hybrid acoustic features, they should elicit larger N200 and smaller P300 components than typical sounds.

Methods

Subjects

A total of 25 participants were tested in this experiment that lasted for about 1 hour. Three participants were excluded from final data analysis because of the high number of trials contaminated by ocular and muscular artifacts. The remaining 22 participants (11 women and 11 men; 19–35 years old) were all right-handed, were nonmusicians (no formal musical training), had normal audition, and had no known neurological disorders. They all gave written consent to participate to the experiment and were paid for their participation.

Stimuli

We first recorded sounds by impacting everyday life objects made of different materials (i.e., wooden beams, metallic plates, and various glass bowls) to insure the generation of realistic familiar sounds. Then, we used a simplified version of the model described in Aramaki and Kronland-Martinet (2006) on the basis of an additive synthesis technique to resynthesize these recorded sounds (44.1-kHz sampling frequency). From a physical point of view, because the vibrations of an impacted object (under free oscillations) can generally be written as a sum of exponentially damped sinusoids, the recorded sounds are considered to be well described by

$$s(t) = \theta(t) \sum_{m=1}^{M} A_m \sin(2\pi f_m t) e^{-\alpha_m t} \qquad (1)$$

where $\theta(t)$ is the Heaviside unit step function, M is the number of sinusoidal components, and the parameters $A_m, f_m,$ and α_m are the amplitude, frequency, and damping coefficient of the mth component, respectively. The synthesis parameters of this model (i.e., $M, A_m, f_m,$ and α_m) were estimated from the analysis of the recorded sounds (examples of analysis–synthesis processes can be found in Kronland-Martinet, Guillemain, & Ystad, 1997). In practice, many sounds from each material category were recorded and resynthesized and the five most representative sounds per category (as judged by seven listeners) were selected for the current study.

All sounds were tuned to the same chroma (note C that was closest to the original pitch) and were equalized in loudness by gain adjustments. Averaged sound duration was 744 msec for wood, 1667 msec for metal, and 901 msec for glass category. Because the damping is frequency dependent (see Equation 1), the damping coefficient of each tuned component was modified according to a damping law estimated from the original sound (Aramaki, Baillères, Brancheriau, Kronland-Martinet, & Ystad, 2007).

A total of 15 sound continua were created as progressive transitions between two material categories (i.e., 5 different continua for each transition: wood–metal, wood–glass, and glass–metal). Each continuum comprised 20 sounds that were generated using additive synthesis (see Equation 1). Each sound of the continuum was obtained by combining the spectral components of the two extreme sounds and by varying the amplitude and damping coefficients. Amplitude variations were obtained by applying a cross-fade technique between the two extreme sounds. Damping coefficients were estimated by a hybrid damping law resulting from the interpolation between the damping laws of the two extreme sounds. Note that this manipulation allowed creating hybrid sounds that differed from a simple mix between the two extreme sounds because at each step, the spectral components are damped following a same hybrid damping law. All sounds are available at http://www.lma.cnrs-mrs.fr/~kronland/Categorization/sounds.html.

Procedure

A total of 300 sounds were presented in a random order within five blocks of 60 sounds through one loudspeaker (Tannoy S800) located 1 m in front of the participants. They were asked to listen to each sound and to categorize it as wood, metal, or glass as quickly as possible by pressing one button on a three-button response box. The association between response buttons and sounds was balanced across participants. The experiment was conducted in a faradized room. A row of XXXX was presented on the screen 2500 msec after sound onset for 1500 msec to give participants time to blink, and the next sound was then presented after a 500-msec delay.

Recording ERPs

The EEG was recorded continuously from 32 Biosemi Pin-type active electrodes (Amsterdam University) mounted on an elastic headcap and located at standard left and right hemisphere positions over frontal, central, parietal, occipital, and temporal areas (international extended 10/20 system; Jasper, 1958): Fz, Cz, Pz, Oz, Fp1, Fp2, AF3, AF4, F7, F8, F3, F4, Fc5, Fc6, Fc1, Fc2, T7, T8, C3, C4, Cp1, Cp2, Cp5, Cp6, P3, P4, PO3, PO4, P7, P8, O1, and O2. Moreover, to detect horizontal eye movements and blinks, the EOG was recorded from flat-type active electrodes placed 1 cm to the left and right of the external canthi and from an electrode beneath the right eye. Two additional electrodes were placed on the left and right mastoids. EEG was recorded at 512-Hz sampling frequency using Biosemi amplifier. The EEG was rereferenced off-line to the algebraic average of the left and right mastoids and filtered with a band-pass of 0–40 Hz.

Data were analyzed using Brain Vision Analyzer software (Brain Products, Munich), segmented in single trials of 2500 msec starting 200 msec before the onset of the sound and averaged as a function of the type of sound (i.e., typical vs. ambiguous).

Results

Behavioral Data

Participants' responses and RTs were collected for each sound and were averaged across participants. Sounds categorized within a category (wood, metal, or glass) by more than 70% of the participants were considered as typical sounds; sounds categorized within a category by less than 70% of the participants were considered as ambiguous sounds. As can be seen in Figure 1 (top), participants' responses are consistent with the position of the sound on the continua so that typical sounds are located at extreme positions and ambiguous sounds at intermediate positions on the continua.

RTs to typical and ambiguous sounds were submitted to repeated measures ANOVAs (for this and following sta-

tistical analyses, effects were considered significant if the p value was equal to or less than .05) that included Type of Sounds (typical vs. ambiguous) and Continua (wood–metal, wood–glass, and glass–metal) as within-subject factors. For typical sounds, only RTs associated to correct responses were taken into account. As shown in Figure 1 (bottom), RTs to typical sounds (984 msec) were shorter than RTs to ambiguous sounds (1165 msec), $F(1, 21) = 74.00, p < .001$. Moreover, the Type of Sounds × Continua interaction was significant, $F(2, 42) = 22.24, p < .001$. Results of post hoc comparisons (Tukey tests) showed that although RTs were shorter for typical than for ambiguous sounds for each continuum ($p < .01$), this difference was larger for the wood–glass continuum (281 msec) than for the wood–metal (184 msec) and glass–metal (79 msec) continua.

Electrophysiological Data

Separate ANOVAs were conducted for midline and lateral electrodes. Type of sounds[1] (typical vs. ambiguous) and Electrodes (Fz, Cz, Pz) were used as factors for midline analyses. Type of Sounds, Hemispheres (left vs. right), ROIs (fronto-central R1, centro-temporal R2, and centro-parietal R3), and Electrodes (three for each ROI: [AF3, F3, FC5]/[AF4, F4, FC6]; [T7, C3, CP5]/[T8, C4, CP6]; and [P7, P3, CP1]/[P8, P4, CP2]) were included for lateral analyses. On the basis of visual inspection of the ERP traces (Figure 2) and results of successive analyses in 50-msec latency windows, the following time windows were chosen for statistical analysis:[2] 0–250 msec (P100–N100–P200),
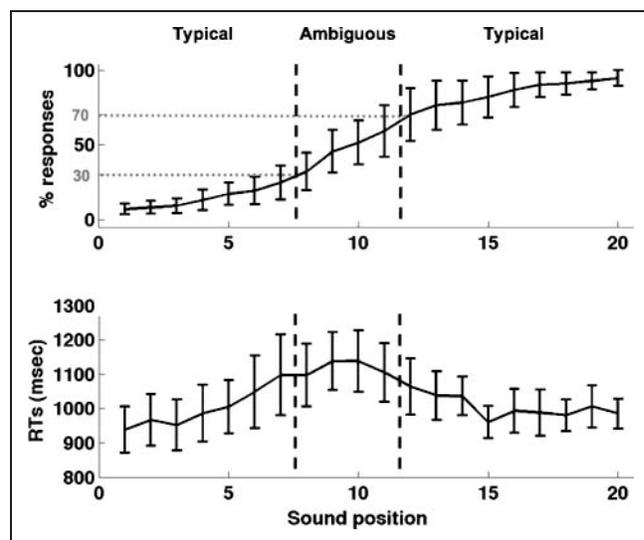
Figure 1. Categorization function (top) and mean RTs in millisecond (bottom) averaged across the 15 continua as a function of sound position (from 1 to 20). Standard deviations are indicated at each sound position. The categorization function represents the percentage of responses in the material category corresponding to the right extreme of the continuum. The vertical dotted gray lines delimit the zones of typical (extremes positions) and of ambiguous (intermediate positions) sounds.

Figure 2. ERPs to typical (black line) and ambiguous sounds (gray line) at midline and at selected lateral electrodes (the most representative electrodes for each ROI). For this and following figures, the amplitude of the effects is represented on the ordinate (in µV; negativity is up). The time from sound onset is on the abscissa (in msec). The gray zones indicate the latency ranges in which differences between typical and ambiguous sounds were significant. These differences started earlier (vertical gray arrows) and were larger (striped zones) over the left than the right hemisphere.
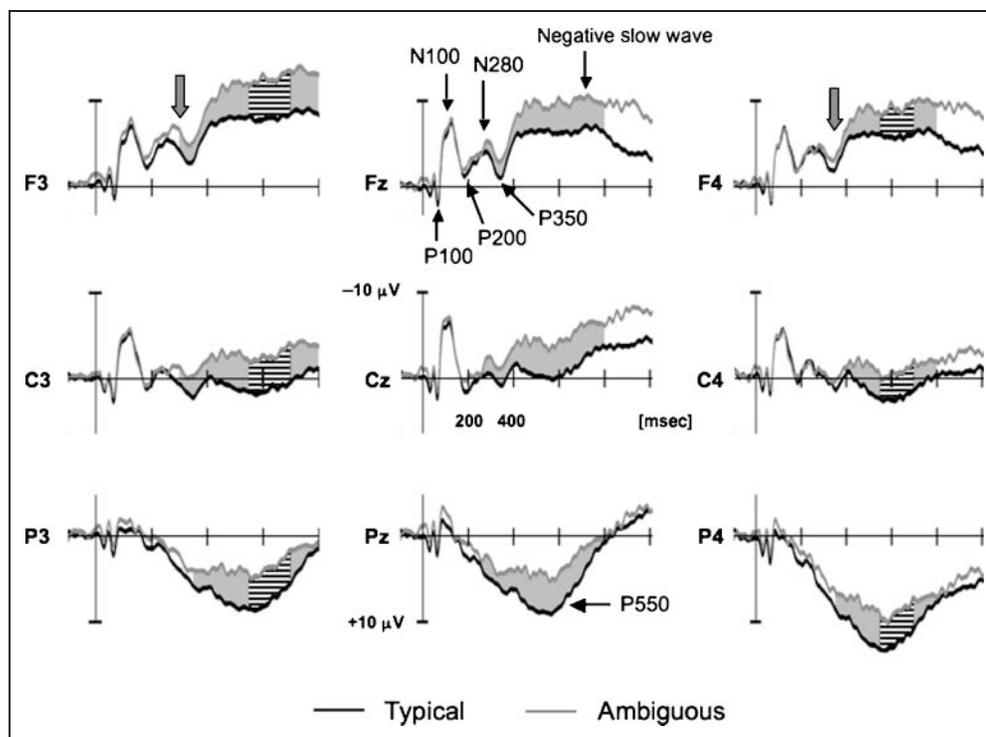
250–400 msec (N280-P350), and 400–800 msec (negative slow wave [NSW] and P550).

Figure 2 shows the ERPs for midline and selected lateral electrodes (the most representative electrode for each ROI). Both typical and ambiguous sounds elicited similar P100, N100, and P200 components at midline and lateral electrodes (no significant effect in the 0- to 250-msec latency band). The ERPs in the two conditions then start to diverge with larger N280 and smaller P350 components in the 250- to 400-msec latency band for ambiguous (-0.89 µV) than for typical sounds (0.54 µV) at midline electrodes, $F(1, 21) = 4.19, p < .05$. Results of fine-grained analyses in successive 50-msec latency bands revealed that these differences started earlier over fronto-central regions of the left—from 300 msec at F3 ($p < .001$) and C3 ($p < .01$) electrodes; Type of Sounds × Hemispheres × ROI × Electrodes interaction: $F(4, 84) = 2.68, p < .05$—than of the right hemisphere—from 350 msec at F4 electrode ($p < .05$); Type of Sounds × Hemispheres × ROI × Electrodes interaction: $F(4, 84) = 2.32, p = .06$ (see Figure 2).

In the 400- to 800-msec latency range, the main effect of Type of Sounds was still significant: midline, $F(1, 21) = 23.60, p < .001$; lateral, $F(1, 21) = 17.91, p < .001$. Although typical sounds were associated with larger positivity (P550) than ambiguous sounds over parietal regions, ambiguous sounds elicited larger NSW than typical sounds over fronto-central regions: Type of Sounds × ROI interaction, $F(2, 42) = 4.11, p < .05$. These differences were larger over the left (3.4 µV) than the right (2.78 µV) hemisphere in the 550- to 700-msec latency band[2]: Type of sounds × Hemispheres interaction, $F(1, 21) = 4.60, p < .05$ (see striped zones in Figure 2).

Discussion

Analysis of behavioral data showed that sounds categorized within a material category by less than 70% of the participants (ambiguous sounds) were associated with slower RTs than sounds that were categorized within a category by more than 70% of the participants (typical sounds). This was found for each continuum. Thus, as hypothesized, ambiguous sounds were more difficult to categorize than typical sounds. This result is in line with previous findings in the literature showing slower RTs for nonmeaningful than for meaningful sounds (e.g., Cummings et al., 2006). The differences between typical and ambiguous sounds were smaller in the wood–metal and glass–metal continua than in the wood–glass continuum. This is interesting from an acoustic perspective because metal sounds typically present higher spectral complexity (related to the density and repartition of spectral components) than both wood and glass sounds that show closer sound properties. Thus, ambiguous sounds in wood–metal and glass–metal continua were easier to categorize than those in the wood–glass continuum and the ambiguity effect was smaller.

Electrophysiological data showed that ambiguous sounds elicited more negative ERPs (a negative component, N280, followed by an NSW) than typical sounds over fronto-central regions. By contrast, typical sounds elicited more positive ERPs (P350 and P550 components) than ambiguous sounds over frontal and parietal regions. These findings were expected on the basis of previous results in categorization tasks showing that the amplitude of the N200 component is larger and the amplitude of

the P300 component is smaller to stimuli that are more difficult to categorize (Donchin & Coles, 1988; Ritter et al., 1983; Duncan-Johnson & Donchin, 1982; Donchin, 1981; Kutas, McCarthy, & Donchin, 1977). Moreover, in line with the long duration of the RTs (around 1 sec), the long latency of the positive component (P550) is taken to reflect the difficulty of the categorization task (participants were required to categorize sounds in one of three possible categories) and the relatively long duration of the sounds (860 msec, on average, over the three categories). Thus, both behavioral and ERP data showed that we were able to create ambiguous sounds that were more difficult to categorize than typical sounds.

The differences between ambiguous and typical sounds started earlier over the left (300 msec) than the right (350 msec) hemisphere and were also larger over the left than right hemisphere in the 550- to 700-msec latency band (see striped zones in Figure 2). This scalp distribution is similar to the left-hemisphere distribution reported for sounds by Van Petten and Rheinfelder (1995). Moreover, as found by these authors, a long-lasting NSW developed over frontal sites that lasted for the entire recording period. Late NSW are typically interpreted as reflecting processes linked with the maintenance of stimuli in working memory, expectancy (Walter, Cooper, Aldridge, McCallum, & Winter, 1964), and attention (King & Kutas, 1995). In the present experiment, the NSW may indeed reflect expectancy processes because a row of XXXX followed sound offset, but it may also reflect sound duration processing (as the "sustained potential" reported by Alain, Schuler, & McDonald, 2002) and categorization difficulty because this fronto-central negativity was larger for ambiguous than for typical sounds. In particular, as it has been proposed for linguistic stimuli (see Kutas & Federmeier, 2000), this larger negativity may reflect the difficulty of accessing information from long-term memory.

Finally, it should be noted that no significant differences were found on the P100 and N100 components. These components are known to be sensitive to sound onset (e.g., attack time) and temporal envelope (for a review, see Kuriki, Kanda, & Hirata, 2006; Shahin, Roberts, Pantev, Trainor, & Ross, 2005; Shahin, Bosnyak, Trainor, & Roberts, 2003; Hyde, 1997). However, because differences in attack time between typical and ambiguous sounds were 0.1 msec, on average, they were consequently not perceptible as this value is below the temporal resolution of the human hearing system (Gordon, 1987), thereby explaining the lack of differences in the ERPs.

EXPERIMENT 2: CONCEPTUAL AND SEMANTIC PRIMING

Results of Experiment 1 showed that we were able to create typical and ambiguous sounds. The goal of Experiment 2 was to use these sounds in a priming design to address the three aims described in the introduction: (1) test for conceptual priming between pairs of nonlinguistic sounds, (2) use only one homogeneous class of sounds (impact sounds), and (3) directly compare conceptual priming for nonlinguistic stimuli on one side and for linguistic stimuli on the other side. To achieve these aims, it was important to use the same task with both types of sounds. Thus, participants were asked to decide whether the target belonged to the same or to a different category than the prime. For the linguistic sounds and based on the design used by Holcomb and Neville (1990), primes were always words, and targets were words, pseudowords, or nonwords (i.e., words played backward). To use similar design and experimental conditions with nonlinguistic sounds, primes were always impact sounds, and targets were typical impact sounds from the same category as the prime, ambiguous sounds and typical impact sounds from a different category than the prime.[3]

On the basis of previous results in the literature with linguistic stimuli (Holcomb & Neville, 1990; Bentin, McCarthy, & Wood, 1985) and results of Experiment 1 with nonlinguistic stimuli, we hypothesized that pseudowords and ambiguous sounds should be more difficult to categorize (i.e., higher error rates and slower RTs) than stimuli from the other two categories. Moreover, as reported by Holcomb and Neville (1990) and in previous studies (for a review, see Kutas, Van Petten, & Kluender, 2006), pseudowords should also elicit larger N400 than words. Holcomb and Neville (1990) argued that "Perhaps this was because their word-like characteristics also produce lexical activation, but because no complete match was achieved, the amount of activation produced was greater and more prolonged" (p. 306). More generally, this result has been taken to reflect the (unsuccessful) search for meaning of orthographically and phonologically legal constructions that nevertheless have no meaning (see Kutas & Federmeier, 2000). However, the N400 to pseudowords may also reflect their lower familiarity than words and their ambiguous nature: They are word-like at the orthographic and phonological levels but are not real words at the semantic level. In such case, ambiguous sounds that share acoustic properties with typical sounds of a material category but nevertheless are not typical exemplars of any categories may also elicit N400-like components. It was therefore of interest to determine whether ambiguous sounds would be processed as pseudowords and elicit N400-like components or would rather elicit increased N280 and NSW as found in Experiment 1. Finally, nonwords (i.e., words played backward) should elicit larger P300 components than words, as reported by Holcomb and Neville (1990). Indeed, although words played backward keep the main attributes of vocal sounds (i.e., the formantic structure of the spectrum due to the resonance of the vocal tract), they should readily be perceived as belonging to a different category than the word prime. Similarly, if typical sounds from a different category than the prime are easily categorized as such, they should also elicit larger P300 components than typical sounds from the same category than the prime.

Methods

Subjects

A total of 19 students (8 women and 11 men; 24 years old, on average) participated in this experiment that lasted for about 1 hour 30 min. None had participated in Experiment 1. They were right-handed, nonmusicians (no formal musical training), French native speakers with no known neurological disorders. They all gave written consent to participate in the experiment and were paid for their participation.

Stimuli

A total of 120 pairs of nonlinguistic stimuli were presented. Primes were always typical sounds from a given material category (i.e., wood, metal, or glass), and targets were either sounds from the same category as the prime (Same condition, 30 pairs), ambiguous sounds (Ambiguous condition, 60 pairs), or sounds from a different category than the prime (Different condition, 30 pairs). The number of ambiguous pairs was twice the number of pairs in the Same and Different conditions to balance the number of Yes and No responses. (On the basis of the results of Experiment 1, we expected participants to give as many Yes as No responses to ambiguous targets.) The averaged duration of nonlinguistic stimuli was 788 msec.

A total of 180 pairs of linguistic sounds were presented. Primes were always French spoken words and targets were spoken words (Same condition, 90 pairs), pseudowords (Ambiguous condition, 45 pairs), or nonwords (Different condition, 45 pairs). Word targets were bisyllabic nouns. Pseudowords were constructed by modifying one vowel from word targets (e.g., _boteau_ from _bateau_). Nonwords were words played backward. The averaged duration of linguistic stimuli was 550 msec.

Procedure

Participants were asked to listen to each pair of stimuli and to determine whether the prime and the target belonged to the same category by pressing one of two response buttons. Nonlinguistic and linguistic stimuli were presented in two separate sessions 10 min apart with the linguistic session always presented after the nonlinguistic session. In each session, pairs of stimuli belonging to the three experimental conditions were randomly presented within three blocks of 40 trials for nonlinguistic pairs and three blocks of 60 trials for linguistic pairs (less nonlinguistic stimuli were presented within a block because sounds were longer in duration than linguistic stimuli).

To balance the number of Yes and No responses, each block of nonlinguistic stimuli comprised 10 same (yes), 20 ambiguous (yes/no), and 10 different pairs (no). Each block of linguistic stimuli comprised 30 Same (yes), 15 Ambiguous (no), and 15 Different (no) pairs. The order of block presentations within the nonlinguistic and linguistic sessions and the association between responses (Yes/No) and buttons (left/right) were balanced across participants.

For both nonlinguistic and linguistic pairs, targets followed prime offset with a 20-msec interstimulus interval. A row of XXXX was presented on the screen 2000 msec after target onset for 2000 msec to give participants time to blink. The prime of the next pair was then presented after a 1000-msec delay.

Recording ERPs

EEG was continuously recorded using the same procedure as in Experiment 1 and later segmented in single trials of 2200 msec starting 200 msec before target onset. Data were analyzed using the Brain Vision Analyzer software (Brain Products, Munich).

Nonlinguistic Sounds

Results

Behavioral data.   For ambiguous sounds, there are no correct or incorrect responses because they can be associated to yes or to no responses. Thus, on the basis of the participants' responses, ANOVAs included Category as a factor with four conditions: Same, Ambiguous/Yes, Ambiguous/No, and Different targets. Results revealed a main effect of Category, $F(3, 54) = 111.71$, $p < .001$: Same and Different targets were associated with low error rates (6% and 4%, respectively) and did not differ from each other ($p = .92$). They differed from Ambiguous/Yes (46%; $p < .001$) and Ambiguous/No targets (53%; $p < .001$) that did not differ from each other ($p = .15$). Mean RTs were not significantly different ($p = .09$: 917 msec for Same, 900 msec for Ambiguous/Yes, 863 msec for Ambiguous/No, and 892 msec for Different targets).

Electrophysiological data.   Two separate ANOVAs were conducted for midline and lateral electrodes. Category (Same, Ambiguous,[4] Different) and Electrodes (Fz, Cz, Pz) were included as factors for midline analyses. Category, Hemispheres (left vs. right), ROIs (fronto-central R1, centro-temporal R2, and centro-parietal R3), and Electrodes (3 for each ROI: [F7, F3, FC1]/[F8, F4, FC2]; [FC5, T7, C3]/[FC6, T8, C4]; and [CP1, CP5, P3]/[CP6, CP2, P4]) were included for lateral analyses. On the basis of visual inspection and results of successive analyses in 50-msec latency windows, time windows chosen for the statistical analysis were 0–150, 150–350, 350–450, 450–550, and 550–700 msec. For Same and Different targets, only correct responses were taken into account. Results are reported in Table 1.

Figure 3 (top) illustrates ERPs to nonlinguistic targets. In all conditions, sounds elicited P100, N100, P200, and

Table 1. Nonlinguistic Targets

(I)	Factors	df	0–150 msec	150–350 msec	350–450 msec	450–550 msec	550–700 msec
Midline	C	2,36	–	–	14.92***	17.13***	12.26***
	C × E	4,72	–	2.60*	2.50*	–	–
Lateral	C	2,36	–	–	15.70***	19.04***	12.10***
	C × ROI	4,72	–	–	2.44*	2.66*	4.89**

(II)	150–350 msec			350–450 msec		
C × E	Fz	Cz	Pz	Fz	Cz	Pz
A − S	–	–	–	−1.74**	−2.24	−2.52
A − D	−1.42*	–	–	−4.23	−3.75	−3.05
S − D	−1.9	–	–	−2.49	−1.51*	–

(III)	350–450 msec			450–550 msec			550–700 msec		
C × ROI	R1	R2	R3	R1	R2	R3	R1	R2	R3
A − S	−1.57	−1.8	−2.55	−1.54	−1.86	−2.37	−1.12**	−1.43	−1.44
A − D	−2.82	−2.51	−3.04	−3.17	−2.97	−4.09	−2.07	−2.04	−3.4
S − D	−1.25	–	–	−1.63	−1.11**	−1.72	−0.95*	–	−1.96

(I) F statistics for the main effect of Category (C), for the Category by Electrodes (C × E) interaction, and for the Category by ROI (C × ROI) interaction in the latency ranges chosen for analyses. (II and III) Mean amplitude differences (in μV) between Same (S), Ambiguous (A), and Different (D) conditions when the C × E and C × ROI interactions were significant. The reported difference values were always significant at $p < .001$ (results of post hoc tests) except when indicated by $p < .05$ or $p < .01$.

*$p < .05$.

**$p < .01$.

***$p < .001$.

N280 components followed by large negative components over fronto-central regions and P550 components over parietal regions. No significant effects were found in the latency range 0–150 msec (see Table 1-I). In the 150- to 350-msec latency range, the main effect of Category was not significant but the Category × Electrodes interaction was significant: Ambiguous and Same targets elicited larger N280 than Different targets at Fz (see Table 1-II). In the 350- to 450-msec latency range, the main effect of Category was significant with larger negativity to Ambiguous than to both Same and Different targets that did not differ from each other except over fronto-central region (see Table 1-II and 1-III). In the 450- to 550-msec latency range, the three conditions significantly differed from each other with the negativity being largest for Ambiguous, intermediate for Same, and the positivity largest for Different targets, with largest differences over centro-parietal regions (i.e., over R3 in Table 1-III). Finally, in the 550- to 700-msec latency range, at midline electrodes, Different targets elicited larger positivity (−0.01 μV) than Same (−2.05 μV) and Ambiguous targets (−3.28 μV) that did not differ from each other. By contrast, at lateral electrodes, the three conditions still differed significantly from each other with largest differences over the centro-parietal region.

Discussion

As hypothesized, the error rate for Same and Different targets was very low (6% and 4%, respectively), which shows that typical sounds were easily categorized as belonging to the same or to a different impact sound category than the prime. By contrast and as expected, on the basis of the results of Experiment 1, ambiguous targets were more difficult to categorize and were categorized as often as belonging to the same (46%) or to a different category (54%) from the prime. This clearly confirms the ambiguous nature of these sounds. The lack of effects on RTs may result from the relatively long duration of the prime sounds (788 msec, on average). Priming effects generally are short lived (Meyer & Schvaneveldt, 1971) and may consequently have vanished by the time the target sound was presented.

Regarding the ERPs, all targets elicited a P100–N100–P200 complex that, as expected (see Discussion of Experiment 1), did not differ between Same, Ambiguous,

Figure 3. ERPs to Same (solid line), Ambiguous (gray line), and Different (dashed line) targets at midline and at selected lateral electrodes (the most representative electrodes for each ROI) for nonlinguistic (top) and linguistic (bottom) stimuli.
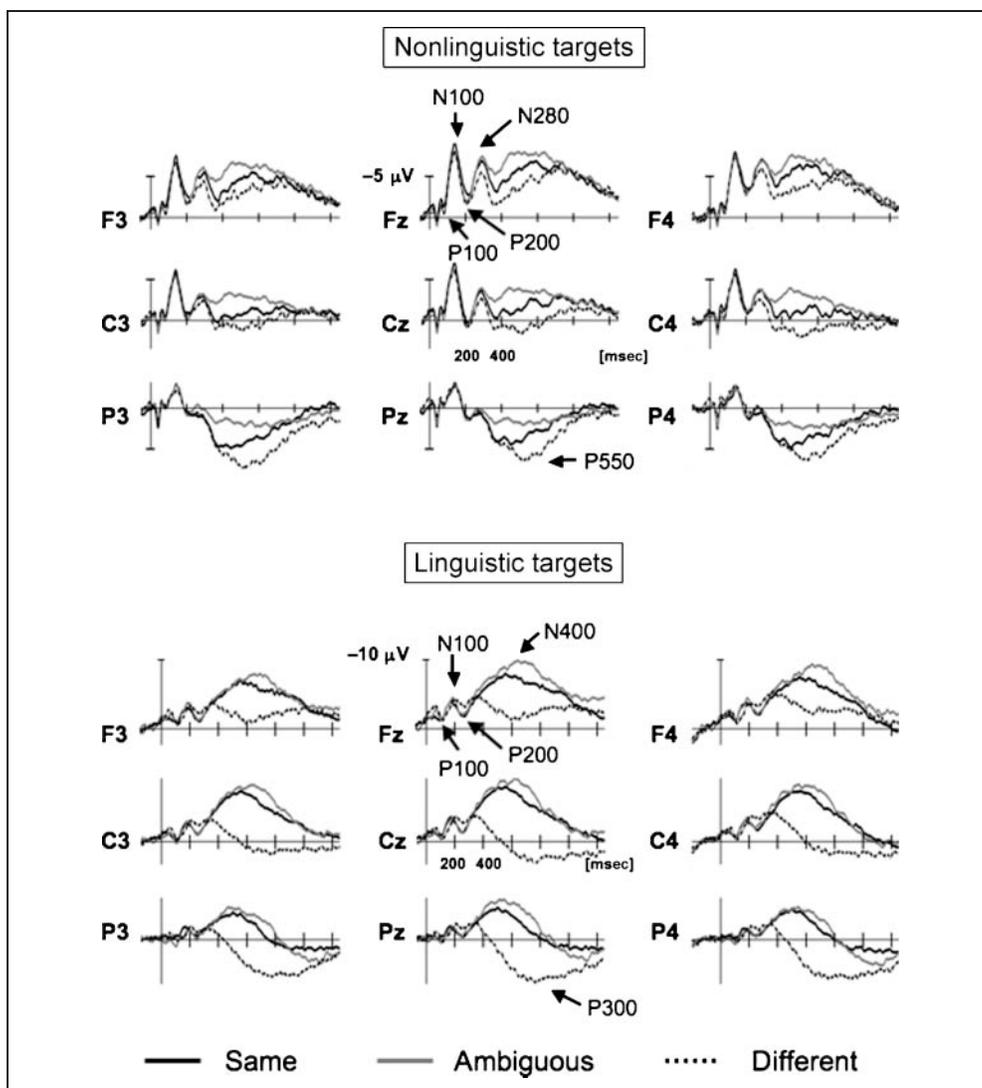
and Different target sounds. An N280 component was also elicited as in Experiment 1. Its amplitude was larger for Same and Ambiguous sounds than for Different sounds over frontal regions. However, the ERPs were morphologically different in Experiments 1 and 2. Although an NSW followed the N280 in Experiment 1 (and lasted until the end of the recording period), a temporally (between 350 and 700 msec) and spatially (fronto-central) localized negativity followed the N280 in Experiment 2. Fine-grained analyses allowed to specify the spatiotemporal dynamics of the effects. First, between 350 and 450 msec, the amplitude of this negative component was largest over fronto-central sites for Ambiguous targets, intermediate for Same targets, and smallest for Different targets. Then, between 450 and 550 msec, typical sounds from a different category than the prime elicited large P300 components over parietal sites thereby reflecting the fact that they were easily categorized as different (4% errors; Holcomb & Neville, 1990; Kutas et al., 1977). Because the same stimuli were used in Experiments 1 and 2, these differences are clearly linked with the task at hand (i.e., in Experiment 1, isolated impact sounds were to be categorized in one of three categories, whereas in Experiment 2, target sounds were compared with a prime). Thus, and as typically shown by fMRI data, these results demonstrate the strong influence of task demands on stimulus processing (e.g., Thierry, Giraud, & Price, 2003).

In Experiment 2, we used a priming design to be able to compare results with previous ones in the literature, and we presented two sounds and no words to reduce the use of linguistic strategies that, as described in the introduction, may have influenced previous results (Orgs et al., 2006, 2007, 2008; Cummings et al., 2006; Plante et al., 2000; Van Petten & Rheinfelder, 1995; and, to a lesser extent, Schön et al., 2009). The finding that a negative component developed in the 350- to 700-msec latency band with largest amplitude to Ambiguous sounds is in line with these previous studies and shows that conceptual priming can occur within sound–sound pairs. Moreover, this result was found when using the homogeneous class of impact sounds. However, before considering the implications of these results for conceptual priming, it is important to

examine results obtained for linguistic targets preceded by linguistic primes.

Linguistic Sounds

Results

Behavioral data. The main effect of Category (word [W], pseudoword [PW], nonword [NW]; within-subject factor) was significant, $F(2, 36) = 48.15$, $p < .001$: The error rate was higher for PW (13%) than for W (4.9%; $p < .001$) and NW (2.3%; $p < .001$). RTs were not significantly dif-ferent ($p = .61$) for PW (1067 msec), W (1057 msec), and NW (1054 msec).

Electrophysiological data. Similar ANOVAs were conducted as for nonlinguistic sounds. Statistical analysis was conducted in the 0–150, 150–350, 350–600, 600–750, and 750–1100 msec latency ranges. Only correct responses were taken into account. Results of statistical analyses are reported in Table 2.

Figure 3 (bottom) illustrates ERPs to linguistic targets. No significant differences were found in the latency ranges 0–150 and 150–350 msec either at midline or at lateral

Table 2. Linguistic Targets

(I)	Factors	df	0–150 msec	150–350 msec	350–600 msec	600–750 msec	750–1100 msec
Midline	C	2,36	–	–	41.32***	90.52***	40.81***
	C × E	4,72	–	–	5.00**	8.97***	11.24***
Lateral	C	2,36	–	–	36.71***	85.53***	54.50***
	C × ROI	4,72	–	–	4.80**	7.41***	7.51***
	C × ROI × H	4,72	–	–	4.93**	2.87*	3.16*

(II)	350–600 msec			600–750 msec			750–1100 msec		
C × E	Fz	Cz	Pz	Fz	Cz	Pz	Fz	Cz	Pz
P − W	–	–	–	−2.57	−2.1	−2.22	−1.22*	–	–
P − N	−4.48	−6.33	−5.68	−7.79	−10.47	−9.58	−3.34	−5.61	−4.61
W − N	−3.53	−5.67	−4.62	−5.22	−8.37	−7.36	−2.12	−5.4	−4.6

(III)	350–600 msec			600–750 msec			750–1100 msec		
C × ROI	R1	R2	R3	R1	R2	R3	R1	R2	R3
P − W	–	–	–	−2.03	−1.57	−1.45	−0.95*	–	–
P − N	−3.03	−4.07	−4.37	−6.08	−7.11	−7.39	−2.58	−3.68	−3.47
W − N	−2.49	−3.65	−3.91	−4.05	−5.54	−5.94	−1.63	−3.25	−3.71

(IV)	350–600 msec		600–750 msec		750–1100 msec	
C × ROI × H	L	R	L	R	L	R
R1	–	−0.68*	−1.72	−2.33	–	−1.43
R2	–	–	−1.52	−1.61	–	–
R3	–	–	−1.61	−1.29	–	–

(I) *F* statistics for the main effect of Category (C), for the Category by Electrodes (C × E) interaction, for the Category by Regions of Interest (C × ROI) interaction, and for the Category by Regions of Interest by Hemispheres (C × ROI × H) interactions in the latency ranges of interest. (II and III) Mean amplitude differences (in μV) between Words (W), Pseudowords (P), and Nonwords (N) conditions for C × E and C × ROI interactions when effects were significant. The reported difference values were always significant at $p < .001$ (results of post hoc tests) except when indicated by $p < .05$. (IV) Mean amplitude differences P − W (in μV) for C × ROI × H interaction. The reported difference values were always significant at $p < .001$ (results of post hoc tests) except when indicated by $p < .05$.

*$p < .05$.

**$p < .01$.

***$p < .001$.

electrodes, but a main effect of Category was found in the 350–600, 600–750, and 750–1100 msec latency ranges at both midline and lateral electrodes (see Table 2-I). In these three latency ranges, NW always elicited larger positivity than both PW and W with largest differences at Cz and over centro-parietal regions (Table 2-II and 2-III). In addition, between 600 and 1100 msec, PW elicited larger negativity than W over right fronto-central regions (Table 2-IV).

Discussion

Behavioral data, showing higher error rate for PW than for both W and NW, are in line with previous results (e.g., Holcomb & Neville, 1990; Bentin et al., 1985). However, no effect was found on RTs, which again may reflect the relatively long duration of stimuli and of RTs (over 1 sec, on average) together with short-lived priming effects (Meyer & Schvaneveldt, 1971). As expected, on the basis of Holcomb and Neville's (1990) results, PW produced larger N400 components than W over anterior sites. Moreover, this N400 effect was larger over the right than the left hemisphere. This "paradoxical lateralization" (Plante et al., 2000, p. 1680) is consistent with previous results showing right-greater-than-left asymmetry of the N400 effect (Kutas et al., 1988; Kutas & Hillyard, 1982). Finally, the rather long latency of this N400 effect is also consistent with the results of Holcomb and Neville (1990), showing that the N400 effect starts earlier and lasts longer in the auditory than in the visual modality. It may also reflect the difficulty to categorize PW that were very similar to words (they were constructed by replacing only one vowel from an existing word). By contrast, NW (i.e., words played backward) was easy to categorize as different from the prime words and elicited a large P300 component with a posterior scalp distribution (Holcomb & Neville, 1990).

Nonlinguistic versus Linguistic Sounds

Because the same design was used for both nonlinguistic and linguistic sounds within the same group of participants, conceptual and semantic priming effects were directly compared by including Stimulus (nonlinguistic vs. linguistic) as a factor.

ANOVAs were conducted in the 350- to 800-msec time window, where significant differences were found for both nonlinguistic and linguistic sounds. Results of statistical analyses are reported in Table 3. The main effect of Stimulus was significant: ERPs to linguistic stimuli were overall more negative than to nonlinguistic stimuli (Table 3-II). Moreover, the main effect of Category was significant with largest N400 to Ambiguous (ambiguous impact sounds and PW), intermediate to Same (same impact sounds and W), and largest positivity to Different targets (different impact sounds and NW) (Table 3-III). Finally, the Stimulus × Category interaction was significant. Although the difference between Ambiguous and Same targets was similar

Table 3. Nonlinguistic versus Linguistic Targets

(I)	Factors	df	F
Midline	Stimulus	1,18	17.48***
	C	2,36	68.57***
	Stimulus × C	2,36	14.43***
Lateral	Stimulus	1,18	24.18***
	C	2,36	65.95***
	Stimulus × C	2,36	12.69***

(II)	Midline	Lateral
Nonlinguistic	−1.52	−0.42
Linguistic	−4.00	−2.98

(III)	Midline	Lateral
S	−3.51	−2.19
A	−4.95	−3.35
D	0.17	0.43

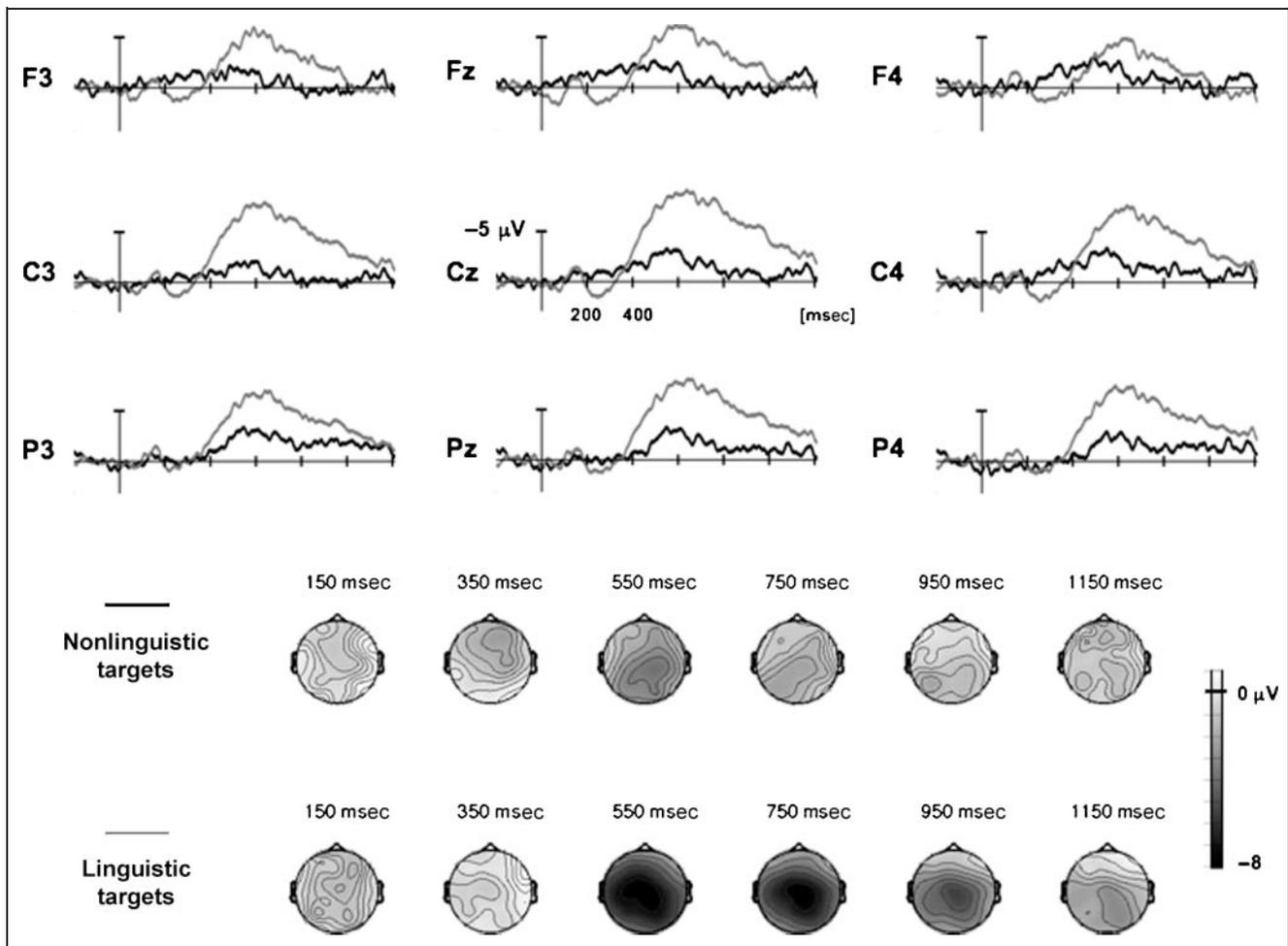| (IV)		Midline	Lateral
A − S	Nonlinguistic	1.43	1.31
	Linguistic	1.46	1.01
D − S	Nonlinguistic	1.80	1.17
	Linguistic	5.57	4.06

(I) *F* statistics for the main effect of Stimulus and Category (C) and for the Stimulus by Category (Stimulus × C) interaction in the 350- to 800-msec latency range. (II) Mean amplitude (in μV) of the main effect of Stimulus. (III) Mean amplitude (in μV) of the main effect of Category: Same (S), Ambiguous (A), and Different (D) conditions. (IV) Mean amplitude differences A − S and D − S (in μV) for Nonlinguistic and Linguistic stimuli. ***$p < .001$.

for both linguistic and nonlinguistic stimuli, the difference between Different and Same targets was significantly larger for linguistic than for nonlinguistic stimuli (Table 3-IV and Figure 4).

GENERAL DISCUSSION

Results of the general ANOVA highlighted clear similarities between conceptual priming for nonlinguistic and linguistic sounds. In both cases, behavioral data showed higher error rates in the Ambiguous than in the Same and Different conditions with no effects on RTs. This ambiguity effect most likely reflects the difficulty to correctly categorize Ambiguous targets as different because they are similar to the prime (e.g., orthographic and phonologic similarity for PW and acoustic proximity for impact sounds). Several studies using priming designs showed higher error rates

Figure 4. Same-minus-Different Difference Waves. ERPs to nonlinguistic (black line) and linguistic (gray line) targets at midline and at selected lateral electrodes (the most representative electrodes for each ROI). Temporal dynamics of the scalp distribution of the effects from 150 to 1150 msec for nonlinguistic and linguistic targets.

for PW than for W (e.g., Holcomb & Neville, 1990) as well as for unrelated than for related words (e.g., Bentin et al., 1985; Boddy, 1981). By contrast, results differ in some studies using nonlinguistic sounds. For instance, Orgs et al. (2006, 2008) found higher error rates in related compared with unrelated pairs. They explained this result by the greater ambiguity of environmental sounds due to causal uncertainties that influence their labeling.

Most interestingly, analyses of the ERPs revealed similar modulation of the late components elicited by nonlinguistic and linguistic sounds: largest negativity for Ambiguous, intermediate for Same, and largest positivity for Different targets. These differences emerged with similar onset latencies in both cases (i.e., at 350 msec after target onset). Importantly, the Stimulus × Category interaction was significant: differences between Same and Different targets were larger for linguistic than for nonlinguistic sounds (see Figure 4). Because linguistic Different targets were words played backward, they were unfamiliar stimuli. Therefore, they were probably more surprising than nonlinguistic Different targets that were typical impact sounds

and consequently more familiar but still different from the prime. By contrast, the priming effect for Ambiguous stimuli was similar in the linguistic and in the nonlinguistic conditions (i.e., the difference between Ambiguous and Same categories was not significantly different, either in amplitude or in scalp distribution for nonlinguistic and for linguistic stimuli).

However, results of separate ANOVAs nevertheless revealed that the spatiotemporal dynamics of the ambiguity effect was somewhat different for nonlinguistic and linguistic sounds, with an earlier onset for ambiguous impact sounds than for PW and a slight predominance over right frontal sites for PW. As noted in the introduction, although priming studies using environmental sounds have reported ERP effects that closely resemble the verbal N400 effect, they also showed differences in scalp distribution. As found here, priming effects were larger over the right hemisphere for words and over the left hemisphere for environmental sounds (Plante et al., 2000; Van Petten & Rheinfelder, 1995). By contrast, Orgs et al. (2006, 2008) found no interhemispheric differences but larger priming

effects for sounds over posterior than anterior sites, and Cummings et al. (2006) found larger differences over anterior than posterior regions (as found here). Thus, the scalp topography seems somewhat variable between experiments, which most likely reflects differences in the acoustic properties of the stimuli and in task demands.

This conclusion is in line with results in the fMRI literature on verbal and environmental sounds showing mixed evidence in favor of the similarity of conceptual priming with nonlinguistic and linguistic sounds. For instance, although both spoken words and environmental sounds activate bilateral temporal regions (Giraud & Price, 2001; Humphries, Willard, Buchsbaum, & Hickok, 2001), Thierry et al. (2003) have demonstrated larger activation of the left anterior and posterior temporal areas for spoken words and larger activation of the right posterior superior temporal areas for environmental sounds. These between-experiments differences were taken to reflect differences in the task semantic requirements. Recently, Steinbeis and Koelsch (2008) provided evidence for both similar and different neural activations related to the processing of meaning in music and speech.

Taken together, our results are in line with previous literature (Schön et al., 2009; Daltrozzo & Schön, 2009; Orgs et al., 2006, 2007, 2008; Cummings et al., 2006; Plante et al., 2000; Van Petten & Rheinfelder, 1995) and argue in favor of the similarity of conceptual priming for nonlinguistic and linguistic sounds. Interestingly, the present results extend previous ones in several aspects. Most importantly, previous results were problematic in that words were always included in the design. As a consequence, the reported conceptual priming effects were possibly due to a linguistic strategy of generating words when listening to sounds. Although we also used linguistic stimuli to be able to compare priming effects within subjects, they were always presented in a separate session. The finding of N400-like components in a sound–sound design, as used in Experiment 2, shows that linguistic mediation is not necessary for an N400-like component to be elicited. Thus, this component may reflect a search for meaning that is not restricted to linguistic meaning. This interpretation is in agreement with the idea that variations in N400 amplitude are related to the "ease or difficulty of retrieving stored conceptual knowledge associated with a word or other meaningful stimuli" (Kutas et al., 2006, p. 10). Moreover, although two conditions (related vs. unrelated) were used in most previous studies, we used three conditions (Same vs. Ambiguous vs. Different) to more closely examine conceptual priming effects. In line with early studies of category membership effects (Ritter et al., 1983; Vaughan, Sherif, O'Sullivan, Herrmann, & Weldon, 1982; Boddy, 1981; Boddy & Weinberg, 1981), stimuli that clearly did not belong to the prime category elicited late positivity (P300 components), whereas stimuli that were ambiguous elicited late negativity (N400-like components) compared with stimuli that belonged to the prime category. Most importantly for our purposes, we were able to demonstrate similar relationships between categories for both non-linguistic and linguistic target sounds.

Conclusion

These results add interesting information to the vast and still largely unexplored domain of the semiotics of sounds. Other experiments using different tasks and stimuli are needed to further explore the similarities and differences in conceptual priming for nonlinguistic and linguistic sounds. However, by using a homogeneous class of environmental sounds (impact sounds), by varying the relationship between prime and target sounds, and by comparing conceptual priming for nonlinguistic and for linguistic sounds within the same participants, we were able to make one step further and to show that conceptual priming develops in a sound–sound design without words and, consequently, that conceptual priming can develop without (or with reduced) linguistic mediation.

Notes

1. The Continua factor was not taken into account to keep enough trials in each condition.
2. Fine-grained analyses were computed as separated ANOVAs (that included the same factors as described in the Results section) in successive 50-msec latency windows from 0 to 800 msec after sound onset. Then, the 50-msec latency windows within which statistically similar effects were found were grouped together between 0–250, 250–400, and 400–800 msec or more specifically between 550 and 700 msec, and an ANOVA was conducted in each latency band.
3. To increase the similarities between the Different conditions for nonlinguistic and linguistic sounds, we also considered the possibility of playing impact sounds backward as was done for the words. However, although such sounds conserve the spectral characteristics of the original sound (i.e., acoustic cues characterizing the material category), they do no longer sound as impact sounds (i.e., the perception of impact disappears). They are therefore ambiguous and difficult to categorize. Because it was important to equate task difficulty for nonlinguistic and linguistic sounds (the words played backward are easy to categorize as different from the prime) and because words played backward keep the main attributes of vocal sounds, we decided to use typical sounds from another material category that are not ambiguous and easy to categorize as different.
4. On the basis of the results of behavioral data showing no differences between ambiguous targets associated with yes and no responses, we averaged ERPs in these two categories together to increase the signal to noise ratio. Moreover, no differences were found when ERPs to Ambiguous/Yes and Ambiguous/No were averaged separately.

REFERENCES

Alain, C., Schuler, B. M., & McDonald, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *Journal of the Acoustical Society of America, 111,* 990–995.

Aramaki, M., Baillères, H., Brancheriau, L., Kronland-Martinet, R., & Ystad, S. (2007). Sound quality assessment of wood for xylophone bars. *Journal of the Acoustical Society of America, 121,* 2407–2420.

Aramaki, M., & Kronland-Martinet, R. (2006). Analysis–synthesis of impact sounds by real-time dynamic filtering. *IEEE Transactions on Audio, Speech, and Language Processing, 14,* 695–705.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 250–267.

Ballas, J. A., & Howard, J. H., Jr. (1987). Interpreting the language of environmental sounds. *Environment and Behavior, 19,* 91–114.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision, and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60,* 343–355.

Boddy, J. (1981). Evoked potentials and the dynamics of language processing. *Biological Psychology, 13,* 125–140.

Boddy, J., & Weinberg, H. (1981). Brain potentials, perceptual mechanisms and semantic categorization. *Biological Psychology, 12,* 43–61.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115,* 92–107.

Daltrozzo, J., & Schön, D. (2009). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience, 21,* 1882–1892.

de Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.

Donchin, E. (1981). Surprise!…surprise? *Psychophysiology, 18,* 493–513.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11,* 357–374.

Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology, 14,* 1–52.

Frey, A., Marie, C., Prod'Homme, L., Timsit-Berthier, M., Schön, D., & Besson, M. (2009). Temporal semiotic units as minimal meaningful units in music? An electrophysiological approach. *Music Perception, 26,* 247–256.

Giraud, A. L., & Price, C. J. (2001). The constraints functional neuroimaging places on classical models of auditory word processing. *Journal of Cognitive Neuroscience, 13,* 754–765.

Gordon, J. W. (1987). The perceptual attack time of musical tones. *Journal of the Acoustical Society of America, 82,* 88–105.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5,* 281–312.

Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *NeuroReport, 12,* 1749–1752.

Hyde, M. (1997). The N1 response and its applications. *Audiology & Neuro-otology, 2,* 281–307.

Jasper, H. H. (1958). The ten–twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

King, J., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7,* 376–395.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7,* 302–307.

Kronland-Martinet, R., Guillemain, P., & Ystad, S. (1997). Modelling of natural sounds by time-frequency and wavelet representations. *Organised Sound, 2,* 179–191.

Kuriki, S., Kanda, S., & Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *Journal of Neuroscience, 26,* 4046–4053.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension, Language comprehension and the N400. *Trends in Cognitive Sciences, 4,* 463–470.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–204.

Kutas, M., & Hillyard, S. A. (1982). The lateral distribution of event-related potentials during sentence processing. *Neuropsychologia, 20,* 579–590.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science, 197,* 792–795.

Kutas, M., Van Petten, C., & Besson, M. (1988). Event-related potential asymmetries during the reading of sentences. *Electroencephalography and Clinical Neurophysiology, 69,* 218–233.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Handbook of psycholinguistics. In M. A. Gernsbacher & M. Traxler (Eds.), *Psycholinguistics electrified II (1994–2005)* (2nd ed., pp. 659–724). New York: Elsevier Press.

Lebrun, N., Clochon, P., Etévenon, P., Lambert, J., Baron, J. C., & 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,* 235–248.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–234.

Orgs, G., Lange, K., Dombrowski, J., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition, 62,* 267–272.

Orgs, G., Lange, K., Dombrowski, J., & Heil, M. (2007). Is conceptual priming for environmental sounds obligatory? *International Journal of Psychophysiology, 65,* 162–166.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2008). N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neuroscience Letters, 436,* 133–137.

Plante, E., Van Petten, C., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal semantic processing in learning disabled adults. *Neuropsychologia, 38,* 1669–1684.

Ritter, W., Simson, R., & Vaughan, H. G. (1983). Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. *Psychophysiology, 20,* 168–179.

Ritter, W., Simson, R., Vaughan, H. G., & Friedman, D. (1979). A brain event related to the making of sensory discrimination. *Science, 203,* 1358–1361.

Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2009). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience, 22,* 1026–1035.

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience, 23,* 5545–5552.

Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., & Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *NeuroReport, 16,* 1781–1785.

Simson, R., Vaughan, H. G., & Ritter, W. (1977). The scalp topography of potentials in auditory and visual discrimination tasks. *Electroencephalography and Clinical Neurophysiology, 42,* 528–535.

Steinbeis, N., & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and fMRI provides evidence for similar and distinct neural representations. *PLoS ONE, 3,* e2226. doi:10.1371/journal.pone.0002226.

Thierry, G., Giraud, A.-L., & Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron, 38,* 499–506.

Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33,* 485–508.

Vaughan, J., Sherif, K., O'Sullivan, R. L., Herrmann, D. J., & Weldon, D. A. (1982). Cortical evoked responses to synonyms and antonyms. *Memory and Cognition, 10,* 225–231.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electrical sign of sensorimotor association and expectancy in the human brain. *Nature, 230,* 380–384.