Psychophysiology, 48 (2011), 645–655. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2010.01134.x

Can out-of-context musical sounds convey meaning? An ERP study on the processing of meaning in music

JULIA GRIESER PAINTER[a] AND STEFAN KOELSCH[a,b]
[a]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b]Cluster of Excellence "Languages of Emotion", Freie Universität Berlin, Berlin, Germany

Abstract

There has been much debate over whether music can convey extra-musical meaning. The experiments presented here investigated whether low level musical features, specifically the timbre of a sound, have a direct access route to meaningful representations. Short musical sounds with varying timbres were investigated with regard to their ability to elicit meaningful associations, and the neural mechanisms underlying the meaningful processing of sounds were compared to those underlying the semantic processing of words. Two EEG experiments were carried out, and N400 effects were found for sound and word targets following sound and word primes in a semantic relatedness judgment task. No N400 effects were found in a memory task. The results show that even short musical sounds outside of a musical context are capable of conveying meaning information, but that sounds require more elaborate processing than other kinds of meaningful stimuli.

Descriptors: Music, Language, Meaning, N400, EEG, ERPs

Music and language have many comparable aspects, and successful efforts have been made to apply linguistic concepts, such as syntax, to music (e.g., Koelsch, Gunter, Wittforth, & Sammler, 2005). With regard to meaning, however, drawing analogies between language and music seems to be more complex. Indeed, whether meaning in music exists and how this may be defined as compared to linguistic semantics has been a matter of much debate (see, e.g., Meyer, 1956; Nussbaum, 2007; Patel, 2008; Raffman, 1992, 1993). In the extensive theoretical discussion on this topic, two kinds of musical meaning are being discussed. They are sometimes referred to as intra-musical meaning and extra-musical meaning (e.g., Patel, 2008).

The former kind of meaning in a musical piece arises when a musical event—such as a tone, a phrase, or a whole section—points to another musical event and leads, for example, to the expectation of that event. Such musical events do not have intrinsic meaning; rather, they become meaningful only through pointing to or implying something else, such as another musical event (e.g., Meyer, 1956). The latter kind of meaning in music arises through the communication of meaning by referring to extra-musical concepts, ideas and events (e.g., Koelsch et al., 2004; Patel, 2008). However, the mechanisms through which extra-musical meaning in music may arise, and how the emergence of musical meaning relates to meaningful representations individuals build in other domains, are largely unknown. Koelsch and Siebel (2005) proposed a model of music perception and processing, which includes the description of neural processes underlying music perception and processing; this model proposes several mechanisms through which meaning in music may arise. It divides perception of music into early and late processing stages. The former include, e.g., feature extraction, Gestalt formation, and the analysis of intervals, whereas the latter include structure building, structural reanalysis, vitalization, and action-related processes. Meaningful representations (as well as emotional responses) may arise from any of these processing stages (with the emotional perception of music also being capable of evoking meaningful representations).

Surprisingly, despite the ongoing theoretical debate regarding the existence of meaning in music (as mentioned above), very little research has been carried out to address the question of whether music can activate meaningful representations. Even fewer studies were aimed at finding neural correlates of the processing of meaning in music. There are, however, empirical studies that support the claim that music can activate meaningful representations by evoking emotions (i.e., Sollberger, Reber, & Eckstein, 2003; Steinbeis & Koelsch, 2008), and thus meaningful concepts related to these emotions. Others have investigated meaningful representations activated by music independent of emotion (e.g., Koelsch et al., 2004). Most of the studies mentioned above used a semantic or affective priming paradigm and could show a behavioral priming effect (Sollberger et al., 2003) or an N400 effect (Daltrozzo & Schoen, 2009a, b; Koelsch et al., 2004; Steinbeis & Koelsch, 2008; Steinbeis & Koelsch, in press).

The authors would like to thank Petr Janata and Jason Golubock for their support in creating the stimuli and Heike Boethel and Ina Koch for their help with collecting data.

Address correspondence to: Dr. Stefan Koelsch, Cluster of Excellence "Languages of Emotion", Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: koelsch@cbs.mpg.de

The N400 is an electrophysiological response that has been shown to be sensitive to the processing of meaningful concepts (e.g., Kutas & Hillyard, 1980). The more difficult a meaningful stimulus is to integrate into a preceding meaningful context, the larger the N400 amplitude will be; this is often referred to as the N400 priming effect. The N400 has mostly been associated with the processing of visually and auditorily presented words and sentences (Kutas & Hillyard, 1980). However, an N400 effect, and thus the processing of meaning related to particular stimuli, has also been found for a wide variety of other material, such as pictures (Hamm, Johnson, & Kirk, 2002; McPherson & Holcomb, 1999), environmental sounds (Cummings et al., 2006; Orgs, Lange, Dombrowski, & Heil, 2006, 2007; van Petten & Rheinfelder, 1995), odors (Grigor, VanTolle, Behan, & Richardson, 1999) and music (Daltrozzo & Schoen, 2009a, b; Koelsch et al., 2004; Steinbeis & Koelsch, 2008).

One way of communicating extra-musical concepts through music is emotion, and meaning in music has thus far mostly been investigated through the pathway of emotional expression (Koelsch et al., 2004; Sollberger et al., 2003; Steinbeis & Koelsch, 2008; Steinbeis & Koelsch, in press). Steinbeis and Koelsch (2008) and Sollberger et al. (2003) both used an affective priming paradigm to investigate the semantic processing of emotional stimuli, confirming that emotional expression in music is one way by which meaning can be conveyed through music. Sollberger et al. (2003) investigated affective priming for word targets following chords and found shorter reaction times on words for affectively congruent chord-word pairs, as compared with incongruent chord-word pairs. Steinbeis and Koelsch (2008) found an effect in the N400 on both chords following words, and words as targets following chords. There was a more negative N400 for affectively incongruent pairs, thus showing for the first time that musical sounds (i.e., chords) are capable of conveying meaning. Thus, one way of accessing meaningful representations of musical elements is via their emotional valence.

While emotion is one important pathway to meaning in music (Steinbeis & Koelsch, 2008), previous studies showed that it is not the only route (Koelsch et al., 2004). One study found neural correlates of the expression of extra-musical meaning while at the same time controlling for emotional expression (Koelsch et al., 2004). Using a priming paradigm, it was found that target words that were meaningfully unrelated to a previous musical context (a musical excerpt of about 10 s) elicited an N400 component that was similar to the N400 component for words that were meaningfully unrelated to a preceding sentence. Very similar effects were found in two other studies, in which musical excerpts had a duration of only 1 s (Daltrozzo & Schoen, 2009a, b). The results of these studies (Daltrozzo & Schoen, 2009a, b; Koelsch et al., 2004) suggest that music can activate representations of meaningful concepts, and that musical information can systematically influence the semantic processing of words as indicated by the N400. This confirms the hypothesis that musical information can convey extra-musical meaning information. Music can thus activate extra-musical concepts and access semantic memory in a similar fashion to other domains, such as language.

One question that has yet to be answered is which elements of music—such as rhythm, harmony, melody, timbre, etc.—are most important in conveying meaning, and whether only the combination of all of these elements in a piece of music or a single element alone can elicit associations. Timbre seems to be an ideal tool to investigate this question, owing to its multi-dimensional nature. Timbre has also already been shown to be capable of eliciting complex and varied verbal associations (Kendall & Carterette, 1993a, b; Solomon, 1958). Thus, investigating how far a single sound with characteristic timbre can elicit extra-musical associations would shed more light on whether and through which pathways sounds can activate semantic representations.

The following experiment was designed to investigate whether meaningful representations can be activated by short sounds with characteristic timbre; the meaningful representations activated by sounds are also compared to those elicited by words. The aim of the experiment was to investigate whether an N400 effect (indicating a meaningful relationship between prime and target stimuli) will be elicited by target words that are meaningfully related to a prime sound (sound-word pairs) or to a prime word (word-word pairs), as had been shown in similar experiments with musical excerpts (Daltrozzo & Schoen, 2009a, b; Koelsch et al., 2004). We also investigated whether target sounds that are meaningfully related to a prime word (word-sound pairs) or a prime sound (sound-sound pairs) would elicit a similar effect in the N400. We hypothesized that a greater N400 amplitude would be observed for target stimuli following meaningfully unrelated prime stimuli, compared to target stimuli following meaningfully related prime stimuli. In addition, we hypothesized that this effect would be found for all types of prime-target combinations (i.e., word-word, word-sound, sound-word, and sound-sound pairs).

EXPERIMENT 1

Methods

Participants

Twenty-four participants took part in the electroencephalogram (EEG) experiment (half of them females, aged from 21 to 27 years, mean age 24.20). All were native German speakers who had no previous history of neurological diseases and were not taking any psychotropic medication. All participants were right-handed (lateralization quotient >90 according to the Edinburgh Handedness Inventory), reported normal hearing, had normal or corrected-to-normal vision, and received payment for their participation. None of the participants had participated in any of the pilot studies, and none of them had received any professional-level musical training, although several participants had received some musical training at some point in their lives.

Stimuli

The stimuli consisted of 32 adjectives and 32 sounds that had been selected based on the results of a pilot study. In the pilot study, participants rated stimulus pairs (360 per participant) with regard to their relatedness (i.e., "How well do these stimuli 'fit' together?") on an eight-point scale. This rating was carried out for sound-word pairs ($n = 128$), sound-sound pairs ($n = 70$), and word-word pairs ($n = 92$). In the pilot study, no additional ratings were obtained for word-sound pairs because they were considered to be the same as those for sound-word pairs. The ratings for each stimulus pair were averaged across participants. The thirty-two stimulus pairs with the highest average relatedness ratings in the pilot study were selected as related pairs, and the thirty-two stimulus pairs with the lowest average relatedness ratings were selected as unrelated pairs. This was done for each type of stimulus combination, resulting in four types of prime-target pairs (word-word, sound-word, word-sound, sound-sound), with 64 prime-target pairs each and 2 levels of relatedness between prime and target stimulus (related, unrelated). The same 32 sounds and words were used for every type of stimulus combination (i.e., word-word, sound-word, word-sound, sound-sound), with the prime-target pairs for the word-sound and sound-word combination being the same. No stimulus was ever combined with itself.

Ratings ($n = 20$) of the emotional valence of all stimuli were conducted independently of the pilot study. The ratings were made for each stimulus independently (i.e., the stimuli were not judged as pairs). They did not show any systematic difference in the emotional valence ratings between related and unrelated prime-target pairs, thus excluding the possibility that a potential effect of relatedness may be confounded with an effect of congruence of emotional valence.

The final sound stimuli consisted of sounds with different timbres; they were all recorded from a Proteus 2000 synthesizer (E-MU systems, www.emu.com), at MIDI note 64 and a recording duration of 2 s. Fifteen of the timbres resembled musical instruments, while the others did not resemble any natural instruments. The fundamental frequency of 7 sounds was E3, 1 sound was E5, 10 sounds were E4, and 14 sounds did not have a single, identifiable pitch (i.e., they were percussive sounds (3) or more complex sounds). The intensity of all sounds was normalized. The average sound duration was 2.74 s (minimum sound duration was 1.05 s and maximum sound duration was 3.19 s). Word stimuli consisted of German adjectives (e.g., *tense*, *open*, *fast*, *strong*, and *colorful*) presented visually with a presentation time of 2700 ms that corresponded approximately to the average sound duration.

Procedure

Participants received written instruction and had the opportunity to practice the task. They were instructed to judge how well the items of each pair of stimuli (word-word, word-sound, sound-word, or sound-sound) fit together (two-alternative forced choice: "fits," "does not fit"), but were given no further instructions as to the criteria they should use for their judgment. They were told that no right or wrong answers existed, but that they should give the answer that first came to mind. Subjects responded by pressing buttons on a button box with their thumbs. The response hands were counterbalanced between participants.

The stimuli pairs were presented in blocks of 8 pairs of the same stimulus combination (i.e., word-word, word-sound, sound-word, or sound-sound) to avoid confusion about the type of the stimulus about to appear. The stimulus combination was announced before the start of each block. The order of the blocks was counterbalanced between participants.

Participants were instructed to avoid eye blinking as much as possible with the exception of an inter-trial interval of 3000 ms in which they saw a cue (***), and during which they had the opportunity to blink. Thus, one trial consisted of a fixation cross (presented for 500 ms), followed by the prime stimulus, an inter-stimulus interval of 700 ms, and the target stimulus (the duration of the target stimulus was the duration of the sound, or 2700 ms for words). The offset of the target stimulus was immediately followed by a green cross, the cue for participants to make a response, presented for the duration of the participant's response. All trials were presented in random order with the constraint that there was a gap of at least 4 stimuli between identical stimuli, and that the same stimulus combination was not presented twice in one block. The sounds were presented via loudspeakers, positioned to the left and right of the participants. Participants were seated in a chair approximately 120 cm away from the screen.

EEG Measurement

The experiment was conducted in an electrically and acoustically shielded room. The EEG was recorded with Ag-AgCl electrodes from 62 scalp locations of the extended 10–20 system (Jasper, 1958) referenced to left mastoid, and 4 electrodes for vertical and horizontal electro-oculogram. Data were recorded with a PORT-I32/MREFA amplifier (Twente Medical Systems International B.V., Oldenzaal, The Netherlands). The sampling rate was 250 Hz, and an online band-pass filter was set from DC to 100 Hz.

Data Analysis

The EEG data were re-referenced to average reference and filtered off-line with a 25 Hz low-pass filter. Artifacts were eliminated by rejecting the EEG data off-line. EEG data were rejected whenever the standard deviation exceeded 33.75 μV or 25 μV for the electro-oculogram (EOG) channels and all other channels respectively in a 200-ms window (to reject fast strong signal changes, such as in eye blinks) and in an 800-ms window (to reject slow strong signal changes, such as in movement-related drifts). On average, 11.65% of trials were rejected. The baseline was defined as the mean amplitude starting 200 ms before stimulus onset to stimulus onset. EEG data were averaged, time-locked to the target stimulus across trials for each stimulus combination, relatedness, participant, and channel, and the averaged data were averaged across channels according to regions of interest (ROIs) for analyses of variance (ANOVAs) (see Table 1 for ROIs).

Two different sets of analyses were conducted. First, the data were averaged according to the pre-experimentally defined categories of relatedness as described above in the methods section. Second, categories were also defined according to participants' individual judgments of the pairs of stimuli in the EEG experiment. If a participant judged a prime-target pair as fitting together, it was assigned to the related category; if it was judged as not fitting together, it was assigned to the unrelated category. In this analysis, one participant had to be excluded because there was less than 50% of the original number of trials left in at least one of the stimulus combinations.

A time window from 350 ms to 500 ms after target-stimulus onset was chosen for statistical analysis based on visual inspection and previous research on the N400 (e.g., Kutas & Federmeier, 2000; Orgs et al., 2006, 2007; Steinbeis & Koelsch, 2008). Several regions of interest were selected for statistical analysis; see Table 1 for details.

Table 1. *Electrode Channels in Each Region of Interest (ROI)*

ROI	Left	Central	Right
Anterior	FP1, AF3, AF7, F3, F5, F7, FT7, FC5, FC3	FPZ, AFZ, FZ, FCZ	FP2, AF4, AF8, F4, F6, F8, FC4, FC6, FT8
Posterior	P7, PO7, PO3, P5, P3, CP5, CP3, C5, C3	CPZ, PZ, POZ, CZ	P8, P6, P4, PO8, PO4, CP4, CP6, C4, C6

To account for violations of the sphericity assumption, only corrected p-values are reported. The Huynh-Feldt correction was used whenever the Huynh-Feldt epsilon was greater than or equal to 0.75. Whenever the Huynh-Feldt epsilon was smaller than 0.75, the Greenhouse-Geisser correction was used, following the guidelines provided by Huynh and Feldt (1976).

The effect sizes are reported for effects of relatedness only, because this effect reflects our main hypothesis. To indicate effect size, r was calculated using Cohen's d (Cohen, 1988). Cohen's d (Cohen, 1988, 1992) was calculated using means and standard deviations (Cohen, 1992; the pooled standard deviation was calculated according to guidelines provided by Thalheimer and Cook, 2002). An r of 0.10 indicates a small effect size, an r of 0.30 a medium effect size, and an $r \geq 0.50$ a large effect size (Cohen, 1992).

After the statistical evaluation, event-related potentials (ERPs) were filtered for visualization purposes with an 11 Hz low-pass filter to remove low-amplitude, high-frequency noise.

Results

Behavioral Results

The agreement was determined by whether a participant's judgment ('fits'/'does not fit') of the stimulus pairs agreed with the pre-experimentally defined categories of relatedness (related/unrelated). The mean agreement across all stimulus combinations was 77.04% ($SD = 7.89$). Agreement for word-word pairs was 76.95% ($SD = 7.29$), for sound-word pairs 78.91% ($SD = 7.02$), for word-sound pairs 80.80% ($SD = 6.07$), and for sound-sound pairs 71.48% ($SD = 8.13$). No reaction times were analyzed because the task was a delayed response task. A within-subjects ANOVA with the factor stimulus combination (i.e., word-sound, sound-word, sound-sound, and word-word) revealed a significant difference for the participants' agreement on relatedness of stimulus pairs between stimulus combinations ($F(1,28) = 5.08$, $p = .023$).

Electrophysiological Results

Figure 1, panel 1, shows the ERPs elicited by target words following related and unrelated prime sounds. Visual inspection showed a negative wave that was maximal at 350 ms started at approximately 250 ms and lasting until approximately 500 ms after target onset. ERPs to related and unrelated target words started to diverge at approximately 350 ms, with the amplitude for unrelated target words being more negative. The same pattern was observed for word targets following word primes, but the difference between related and unrelated targets appeared to be less pronounced (see Figure 1, panel 2). In both stimulus combinations, the effect was distributed globally across the whole scalp. Sound targets elicited a large negative component starting at approximately 350 ms and peaking at approximately 420 ms, with ERPs to related compared to unrelated targets starting to diverge at roughly 400 ms after target onset (Figure 1,

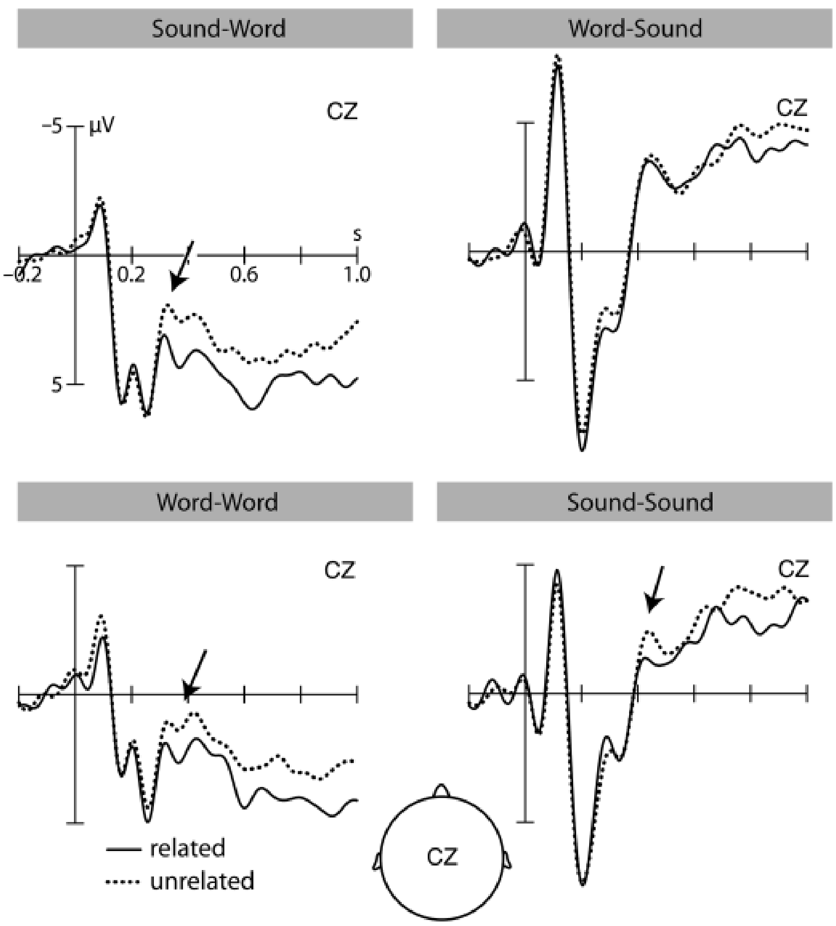

Figure 1. Grand-average ERPs in response to meaningfully related and unrelated target stimuli for pre-experimentally defined categories of relatedness. Unrelated targets elicit a more negative amplitude than related targets in the time window of 350 ms to 500 ms after target onset.

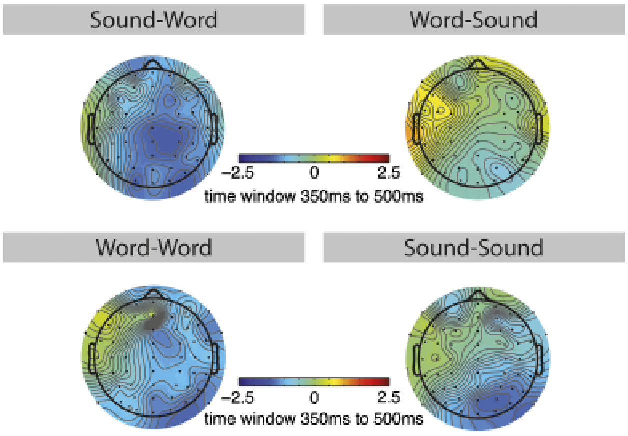

Figure 2. Scalp distribution of the response to unrelated vs. related targets in the time window 350 ms to 500 ms for pre-experimentally defined categories of relatedness. The maps show the difference ERPs (unrelated minus related).

panels 3 and 4). The ERPs to unrelated target sounds showed a more negative amplitude, compared to related targets. The difference between related and unrelated targets appeared to be more distinct in the sound-sound stimulus combination (Figure 1, panel 4). The effect appeared to be globally distributed in both stimulus combinations, but more pronounced over right-posterior electrodes in the word-sound stimulus combination, and more pronounced over frontal electrodes in the sound-sound stimulus combination (see Figure 2, panels 3 and 4).

A 2 × 2 × 3 × 4 MANOVA was carried out to compare all types of stimulus combinations, with the factors stimulus combination (word-word, sound-word, word-sound, and sound-sound), relatedness (related, unrelated), anterior-posterior, and lateralization (left, central, and right ROIs; see Table 1). The analysis revealed main effects of stimulus combination ($F(3,69) = 37.40$, $p < .001$, reflecting that visually presented word targets elicited different ERPs than auditorily presented sound targets), relatedness ($F(1,23) = 4.73$, $p = .04$, $r = 0.10$), reflecting that ERPs elicited by unrelated targets differed from those elicited by related targets, anterior-posterior ($F(1,23) = 221.05$, $p < .001$), and lateralization ($F(2,46) = 6.28$, $p = .009$). Two-way interactions were indicated between factors of relatedness and lateralization ($F(2,46) = 10.33$, $p < .001$), as well as between factors stimulus combination and lateralization ($F(6,138) = 3.08$, $p = .025$). There was no interaction between factors relatedness and stimulus combination ($F(3,69) = 0.88$, $p = .448$, indicating that the amplitude of the N400 effect did not significantly differ between stimulus combinations), and no three- or four-way interactions.

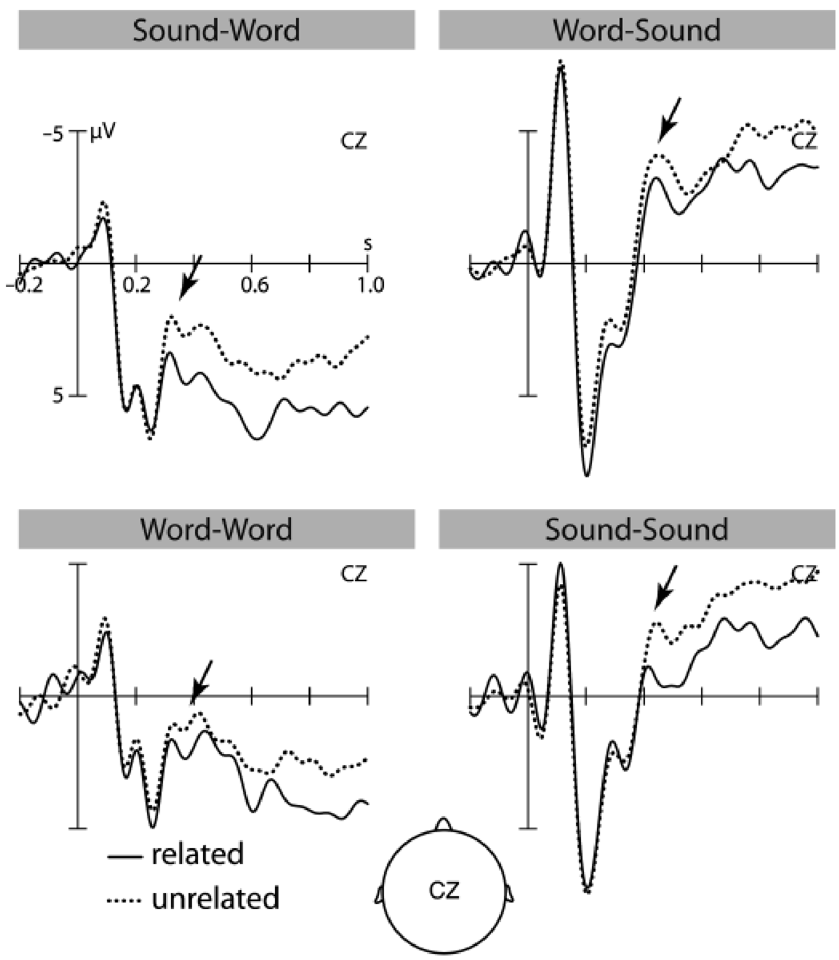

Figure 3. Grand-average ERPs in response to meaningfully related vs. unrelated target stimuli with categories of relatedness defined according to participants' individual judgments. Unrelated targets elicit a more negative amplitude than related targets in the time window of 350 ms to 500 ms after target onset.

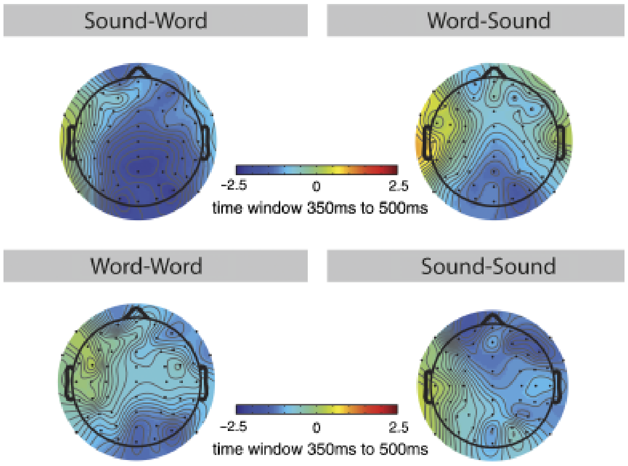

Figure 4. Scalp distribution of the response to unrelated vs. related target stimuli in the time window 350 ms to 500 ms for categories of relatedness defined according to participants' individual judgments. The maps show the difference ERPs (unrelated minus related).

Due to the interaction between lateralization and relatedness (reflecting that the N400 effect elicited by unrelated targets was more pronounced over central and right-hemisphere leads) a follow-up analysis was carried out for central and right ROIs only. This 4 × 2 × 2 MANOVA, with factors stimulus combination, relatedness, and anterior-posterior, revealed main effects of stimulus combination ($F(3,69) = 35.44$, $p < .001$), relatedness ($F(1,23) = 6.41$, $p = .018$, $r = 0.12$), and anterior-posterior ($F(1,23) = 185.27$, $p < .001$), but again no interaction between factors relatedness and stimulus combination, and no other two- or three-way interaction.

Although the latter MANOVA did not yield an interaction between relatedness and stimulus combination, we also compared ERPs to related and unrelated targets for the right and central ROIs separately for each stimulus combinations, to obtain a more detailed impression about the strength of the N400 effects elicited in each stimulus combination. These analyses indicated an effect for sound-word pairs ($F(1,23) = 6.13$, $p = .021$, $r = 0.15$), a marginally significant effect for word-word pairs ($F(1,23) = 3.58$, $p = .07$, $r = 0.14$), and non-significant effects for word-sound pairs ($F(1,23) = 0.25$, $p = .623$, $r = 0.04$), and sound-sound pairs ($F(1,23) = 2.42$, $p = .133$, $r = 0.15$).

The results described above were obtained when the pre-experimentally defined categories of relatedness were used for the analyses (see Methods). In addition, categories were also defined according to participants' individual judgments of the pairs of stimuli in the EEG experiment (i.e., participants' judgment on the relatedness of stimulus pairs). See Figures 3 and 4 for ERPs and scalp distributions of the response to unrelated vs. related target stimuli.

As in the previous analysis, we first computed a 2 × 2 × 3 × 4 MANOVA on the ERPs sorted by the individual relatedness-judgments of participants. This MANOVA included again the factors stimulus combination (word-word, sound-word, word-sound, and sound-sound), relatedness (related, unrelated), anterior-posterior, and lateralization (left, central, and right ROIs). Results revealed main effects of stimulus combination ($F(3,66) = 34.03$, $p < .001$), relatedness ($F(1,22) = 28.10$, $p < .001$, $r = 0.15$), anterior-posterior ($F(1,22) = 209.17$, $p < .001$), and lateralization ($F(2,44) = 6.36$, $p = .010$). Significant two-way interactions were yielded between factors stimulus combination and lateralization ($F(6,132) = 3.07$, $p = .024$), as well as between factors relatedness and lateralization ($F(2,44) = 17.72$, $p = .001$). There were no significant three-way interactions. Due to the interaction between lateralization and relatedness (reflecting that the N400 effect elicited by unrelated targets was more pronounced over central and right-hemisphere leads) a follow-up analysis was carried out for central and right ROIs only. This 4 × 2 × 2 MANOVA, with the factors stimulus combination, relatedness, and anterior-posterior, revealed main effects of stimulus combination ($F(3,66) = 32.18$, $p < .001$), relatedness ($F(1,22) = 36.86$, $p < .001$, $r = 0.17$), and anterior-posterior ($F(1,22) = 176.39$, $p < .001$), but no interactions.

Although the latter MANOVA, as with the analysis using the pre-experimentally defined categories of relatedness, did not yield an interaction between relatedness and stimulus combination, we also compared ERPs to related and unrelated targets for the right and central ROIs separately for each stimulus combinations, to get a more detailed impression about the strength of the N400 effects elicited in each stimulus combination. These analyses indicated an effect for sound-word pairs ($F(1,22) = 11.71$, $p = .002$, $r = 0.20$; see Figure 3, panel 1), and sound-sound pairs ($F(1,22) = 5.15$, $p = .033$, $r = 0.22$; see Figure 3, panel 4), and marginally significant effects for word-word pairs ($F(1,22) = 4.16$, $p = .053$, $r = 0.13$; see Figure 3, panel 2), and word-sound pairs ($F(1,22) = 3.05$, $p = .094$, $r = 0.13$; see Figure 3, panel 3).

To test whether the N400 effects elicited in the two types of relatedness categories (pre-experimentally defined categories of relatedness and relatedness defined according to participants' individual judgments) differed, we computed a MANOVA for the right-central ROI with the factors stimulus combination (word-word, sound-word, word-sound, and sound-sound), relatedness (related, unrelated), anterior-posterior ROIs, and the between-subject factor type of relatedness category (pre-experimentally defined categories of relatedness and relatedness defined according to participants' individual judgments). This MANOVA did not show an effect of relatedness category ($F(1,45) = 0.01$, $p = .923$) and no significant two-way interaction between factors relatedness and type of relatedness category ($F(1,45) = 1.01$, $p = .318$).

Discussion

The results showed an N400 effect for related compared to unrelated target stimuli, independent of the type of stimulus combination, confirming our main hypothesis that sounds with characteristic timbres can elicit meaningful associations. This N400 effect was more pronounced over central and right scalp locations, and there was no significant difference in the N400 effect when relatedness was analyzed according to the pre-experimentally defined categories of relatedness (in the independent behavioral pilot experiments) compared to when the participants' individual judgments during the experiment were considered. When relatedness was analyzed according to the judgments that had been obtained pre-experimentally (in the independent behavioral pilot experiments), a closer look at the effect sizes for each of the four different stimulus combinations (word-word, word-sound, sound-word, and sound-sound), however, revealed that nominally larger effect sizes were found for the word-word, sound-word, and sound-sound stimulus combinations, and smaller ones for the word-sound combination (the effects for the sound-sound and word-sound combinations were not statistically significant).

When looking at the effects for each of the stimulus combinations for the relatedness categories defined according to participants' individual judgments, the effect sizes were larger overall compared to the effects in the analysis of the pre-experimentally defined categories of relatedness. Moreover, even though there was no significant difference between the two types of analysis (pre-experimentally defined categories of relatedness vs. categories defined according to participants' individual judgments), the effect sizes indicate that the overall relatedness effect found with the latter analysis was larger. The judgments indicating whether a participant perceived stimulus pairs as fitting or not fitting together were highly similar across participants. This shows that, while participants generally agreed on the relatedness of stimulus pairs with regard to the meaningful associations they evoked, the difference between pre-experimentally established categories of relatedness and categories established according to participants' judgments in the present experiment seems to be great enough to influence the size of the N400 effect.

EXPERIMENT 2

In Experiment 1, an overall N400 effect was found in all four stimulus combinations. However, the subjects' task was to judge explicitly whether a meaningful relationship between two stimuli existed. Therefore, no conclusions can be made as to whether meaningful representations are automatically activated even when meaningful processing of a stimulus is not required to complete a task. This issue was addressed in Experiment 2, in which we investigated whether differences in the N400 would still be found with a less elaborate processing of the stimuli; that is, with a task that does not require processing of the meaning of a stimulus. The experimental paradigm was identical to Experiment 1, except that a memory task was employed (as in Koelsch et al., 2004): Participants were asked to focus on recognizing the stimuli they had heard or read in the previous experimental block. The participants were thus not required to attend to the meaning of the stimuli and did not have to make a judgment as to the relatedness of the stimulus pairs. Moreover, they were not alerted to, or informed about, a possible connection between stimulus pairs.

Methods

Subjects

Twenty subjects, who were paid for their participation, took part in Experiment 2 (10 females). All subjects were native speakers of German who had no previous history of neurological diseases and were not taking any psychotropic medication. The age range was 18 to 29 years ($M = 23.8$ years). All participants were right-handed (lateralization quotient >90 according to the Edinburgh Handedness Inventory), reported normal hearing, and had normal or corrected-to-normal vision. None of the participants had participated in any of the pilot studies or Experiment 1, and none of them had received any professional-level musical training.

Stimuli and Procedure

Stimuli and procedures were the same as in Experiment 1, except that the instructions and the task varied. Participants were instructed to pay close attention to the stimuli due to a memory task they had to complete after each block. After each block (a block contained eight stimulus pairs), a probe stimulus (either a word or a sound) randomly chosen from all stimuli was presented, and participants had to judge whether that stimulus had been presented in the previous block. They had to push the left button on a button-box if they thought the stimulus had appeared in the last block and the right button if they thought the stimulus had not been presented. Feedback (“correct” or “wrong”) was given after subjects had given the response.

EEG Measurement and Data Analysis

The EEG measurement and data analysis were identical to Experiment 1. On average, 11% of trials were rejected. *P*-values are reported according to the same criteria as in Experiment 1. After the statistical evaluation, ERPs were filtered for visualization purposes with the same filter as in Experiment 1.

Results

Behavioral Results

The average error rate (misses and false alarms) for all stimulus combinations was 19.6% ($SD = 10.4\%$). Table 2 shows hits, false alarms (FA), correct rejections (CR), and misses as a percentage of total responses, for each stimulus combination.

Electrophysiological Results

Figure 5 (panels 1 and 2) shows the ERPs elicited by target words following related and unrelated prime sounds and words. Visual inspection showed that both related and unrelated target words elicited a negative wave that was maximal at 350 ms after stimulus onset, started at approximately 200 ms, and lasted until approximately 500 ms after target onset. No difference was visible in the ERPs of unrelated compared to related targets. Similarly, both related and unrelated sound targets elicited a distinct negative component starting at approximately 200 ms and peaking at around 420 ms (Figure 5, panels 3 and 4), with virtually no difference between ERPs to unrelated and related targets.

As in Experiment 1, a $2 \times 2 \times 3 \times 4$ MANOVA was carried out with the factors stimulus combination (word-word, sound-word, word-sound, and sound-sound), relatedness (related, unrelated), anterior-posterior ROIs, and lateralization (right, middle, central ROIs). The analysis revealed a main effect of anterior-posterior ($F(1,19) = 107.40$, $p < .0001$) but no other main effect and no two-way interactions. There was one significant three-way interaction between factors anterior-posterior, stimulus combination, and lateralization ($F(6,114) = 3.55$, $p < .0045$), but no other significant three-way or four-way interactions. Even when analyzing the ERPs of the right and central ROIs only (analogous to Experiment 1), no effect of relatedness was found ($F(1,19) = 0.38$, $p = .54$).

To test whether the effects found in Experiment 1 differed from those found in Experiment 2 (with the pre-experimentally defined categories of relatedness), we computed a MANOVA with the factors stimulus combination (word-word, sound-word,

Table 2. *Hits, False Alarms (FA), Correct Rejections (CR), and Misses are Reported as Percentages of Total Responses*

Stimulus combination	FA	Hit	Miss	CR
Word-Word	2.63	27.63	9.21	60.52
Word-Sound	8.55	17.10	7.23	67.10
Sound-Word	11.84	15.78	7.89	64.47
Sound-Sound	17.88	31.12	13.24	37.74

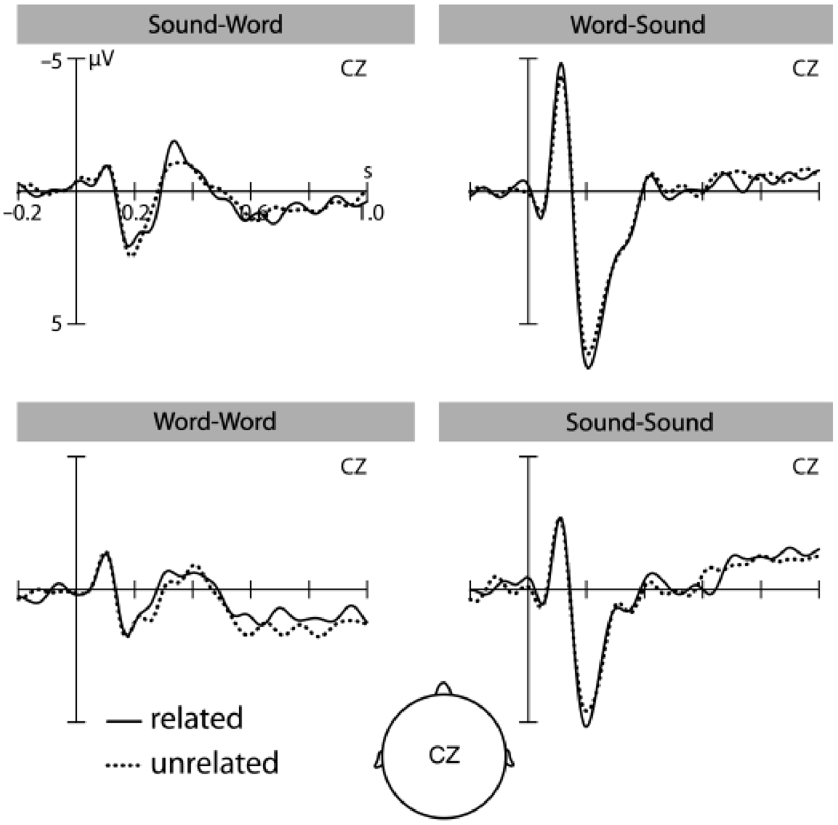

Figure 5. Grand-average ERPs in response to meaningfully related vs. unrelated target stimuli. The ERP waveforms did not differ statistically from each other in the time window from 350 ms to 500 ms after target onset.

word-sound, and sound-sound), relatedness (related, unrelated), anterior-posterior ROIs, lateralization (left, central, and right ROIs), and the between-subjects factor Experiment. This MANOVA did not show an effect of Experiment ($F(1,37) = 0.80$, $p = .378$), but a significant interaction between factors relatedness and experiment ($F(1,37) = 4.08$, $p = .050$, reflecting that an N400 effect was elicited in Experiment 1, but not in Experiment 2), as well as interactions between factors stimulus combination and experiment ($F(3,111) = 22.47$, $p < .001$), and anterior-posterior and experiment ($F(1,37) = 115.72$, $p < .001$). One significant three-way interaction, that of relatedness, lateralization, and experiment was found ($F(2,74) = 5.11$, $p = .008$, due to the N400 being maximal over right and central ROIs in Experiment 1).

Discussion

The ERPs showed no significant differences in amplitude between related and unrelated stimuli in a time window from 350 ms to 500 ms after stimulus onset, and thus no N400 effect, despite the fact that participants did attend to the stimuli and followed the instructions (as indicated by the behavioral data). This indicates that, even though participants attended to the stimuli in order to complete the task, the stimuli were not processed as meaningful information.

A study very similar to the present one could demonstrate an N400 effect with a memory task (Koelsch et al., 2004) and musical excerpts. However, the stimuli in that study (Koelsch et al., 2004) were excerpts of music and sentences that had a longer duration (i.e., 10 s) than the stimuli used in the present experiment. Perhaps the longer duration of stimuli led to a more elaborate processing; however, this issue remains to be elaborated on.

The absence of an N400 effect, even for word-word pairs, supports the hypothesis that the lack of an N400 effect is due to the task. N400 effects have been found for a variety of tasks, such as lexical decision (e.g., Anderson & Holcomb, 1995; Bentin, McCarthy, & Wood, 1985; Daltrozzo & Schoen, 2009b), letter search (Heil, Rolke, & Pecchinenda, 2004; Kutas & Hillyard, 1989), and others (e.g., object identification, McPherson & Holcomb, 1999, and fragment matching, van Petten & Rheinfelder, 1995). The N400 effect generally seems to be smaller when semantic processing is not necessary to carry out a task, such as a pitch comparison task (Hohlfeld & Sommer, 2005) or a memory task (Bentin, Kutas, & Hillyard, 1993). In a recent study with short musical excerpts and words, Daltrozzo and Schoen (2009b) also found a diminished and later-than-usual (i.e., 500 ms to 650 ms after target onset) N400 effect for a lexical decision task compared to a task in which a semantic judgment had to be made. However, some studies also failed to find an N400 effect for non-semantic tasks, such as, e.g., Bentin et al. (1993), who found no N400 effect when using a non-word counting task with an auditory paradigm. Several studies also reported an absence of an N400 effect for a visual paradigm. Chwilla, Brown, and Hagoort (1995) found no N400 in a task that emphasized only physical aspects of the stimuli (i.e., letter size), similar to Deacon, Breton, Ritter, and Vaughan (1991), who also found no N400 effect with a size discrimination task. Taking these disparate findings into consideration, it appears that the presence and size of an

N400 effect can be greatly influenced by the task, which may account for the effect of task found between Experiment 1 and Experiment 2, and the absence of an N400 effect in Experiment 2.

Another factor may be contributing to the lack of activation of semantic representations by words. Words were selected with the primary aim of describing the meaning of the sounds. These attributes were chosen in such a way as to reflect a maximally diverse (i.e., diverse meanings) set of words to describe a wide variety of sounds. While the word-word pairs still resulted in high relatedness ratings (an average of 76.95% agreement on the relatedness/unrelatedness of word-word pairs), it is possible that the associations evoked by those words were not as strong as might be expected with a stimulus set designed purely for a linguistic task. In general, differences in the degree of relatedness of prime and target pairs have been found to lead to a difference in the magnitude of the N400 effect (see Kutas & Federmeier, 2000, for a review). Combined with expected weaker N400 effects due to the nature of the task (as found in other studies), this may have contributed to the lack of an N400 effect found in Experiment 2 for word-word pairs.

General Discussion

The findings of Experiment 1 demonstrate that the perception of a sound, even when presented outside of a musical context, can significantly influence the meaningful processing of a subsequent word or sound (as indexed by the N400). Similarly, the perception of a word can influence the subsequent processing of another word or sound. This shows that single sounds—even when presented outside of a musical context—can activate representations of meaningful concepts.

The results showed that an N400 effect can be observed for all four types of prime-target combinations (word-word, word-sound, sound-word, and sound-sound) with this effect being more pronounced over central and right scalp locations. Experiment 2 was conducted with the same stimulus material as Experiment 1, but with a different task; the analyses did not reveal any effects in the N400 time-window.

The N400 effect found in Experiment 1 complements other studies on extra-musical meaning in music (Koelsch et al., 2004; Steinbeis & Koelsch, 2008; Steinbeis & Koelsch, in press). A similar N400 effect for word targets following musical primes was found by Koelsch et al. (2004); however, they only used word targets and thus could not show an N400 effect for sound targets. In addition, this study shows that even a short sound stimulus presented outside of any musical context, such as a melody or rhythm, can elicit meaningful associations. An N400 effect on a sound target has been observed in other studies (Daltrozzo & Schoen, 2009a, b; Steinbeis & Koelsch, 2008). However, the meaningful context established in the experiment by Steinbeis and Koelsch (2008) was based on affective priming. The studies by Daltrozzo and Schoen (2009a, b) did not use an affective priming paradigm, but also did not control explicitly for the affective values of the prime target pairs they used. Moreover, the stimuli were musical excerpts. In the present study, emotion as a cause of meaning that is conveyed by a sound was ruled out because there was no systematic difference in affective value between the stimulus pairs. In addition, single sounds presented outside of a musical context were used as stimuli.

With regard to N400 effects in other domains, direct comparisons should be drawn with caution because other types of stimuli, such as pictures, words, and even environmental sounds are probably not processed in the same way as musical sounds with regard to their meaningful representations. The N400 effect usually lasts from 250 ms to 500 ms after the onset of a meaningful stimulus and typically peaks at approximately 400 ms (Kutas & Federmeier, 2000). The effect is broadly distributed across the scalp and, even though modality and domain-specific differences in lateralization have been found, the N400 seems to be quite similar across domains and modalities (see, e.g., Cummings et al., 2006; McPherson & Holcomb, 1999; Orgs et al., 2006, 2007; see van Petten & Luka, 2006 for a review regarding auditory and visual N400). The results of the present study thus seem to be in line with the general time window and distribution of the N400 effect found in other studies (e.g., McPherson & Holcomb, 1999; Orgs et al., 2006, 2007).

As mentioned above, similar experiments have been conducted with chords (Steinbeis & Koelsch, 2008), very short musical excerpts (Daltrozzo & Schoen, 2009a) and environmental sounds (Orgs et al., 2006, 2007; van Petten & Rheinfelder, 1995). It seems that these studies come closest to the present study as a basis for comparison in terms of the scalp distribution, as well as the time window of the N400 effect. When comparing the present study to N400 effects found for environmental sounds, the picture is quite varied. Whereas Orgs et al. (2006) found a more frontal distribution for visually presented target words as compared to the environmental sounds, but no lateralization, van Petten and Rheinfelder (1995) reported a more left lateralized N400 for the sound targets as compared to word targets. This is at odds with the findings of the present study, where both target sounds following sounds and sounds following words elicited an evenly distributed effect in the N400, with a stronger effect found for central and right ROIs. Words also elicited an evenly distributed N400, with a stronger effect over central and right ROIs, which would be consistent with the findings of van Petten and Rheinfelder (1995). Steinbeis and Koelsch (2008) report a broadly distributed N400 for target chords, and Daltrozzo and Schoen (2009a) found a bilaterally distributed parietal N400 effect for short musical excerpts and target words; this supports the observation in Experiment 1 of an N400 elicited by sound targets being broadly distributed. However, the effect found in the present study appears to be stronger over central and right scalp locations. With regard to the time window of the N400 effect for environmental sounds, van Petten and Rheinfelder (1995) report a negative going component starting at around 300 ms and extending to 700 ms. Orgs et al. (2006, 2007) find an N400 component between 200 ms and 500 ms post-stimulus for both words and sound targets. The time window of the N400 effect found in Experiment 1 (350 ms to 500 ms) thus seems to be partly overlapping with the time window found for environmental sounds; the differences might be explained by the differences between environmental sounds and the sounds that were used in the present study. The time window found in Experiment 1 (350 ms to 500 ms after target onset) is also quite similar to the time windows described by Steinbeis and Koelsch (2008) and Koelsch et al. (2004), who both found an N400 effect in a time window from 300 ms to 500 ms post stimulus onset. However, the N400 effect found by Daltrozzo and Schoen (2009a), with a semantic judgment task, starts earlier (250 ms) and extends longer (650 ms). Daltrozzo and Schoen (2009b) also report a later time window for an N400 effect (500 ms to 650 ms) with a lexical decision task.

Contrary to findings in other domains and modalities, such as auditorily or visually presented words, pictures, or environmental sounds (Hamm et al., 2002; Kutas & Hillyard, 1980; Orgs et al., 2006, 2007), the task that participants were given here seemed to have a considerable influence on the existence of an N400 effect. Even though an N400 effect has been observed for words following music with a memory task before (Koelsch et al., 2004), no such effect was found in Experiment 2. It therefore seems that the N400 effect found here is sensitive to the level of processing, because an effect was only found in a task that emphasized the meaningful processing of a sound and not in the memory task, which presumably does not require the processing of the meaningful associations of a sound.

In addition, the length of the stimulus, that is, the time window over which a meaningful context is built up, might interact with the task. Daltrozzo and Schoen (2009b) also used short sound stimuli (1 s) with an implicit task; the N400 effect they found, however, was smaller than in a semantic judgment task (Daltrozzo & Schoen, 2009a, b) and occurred much later (500 ms to 600 ms). The task being used was a lexical decision task and not a memory task, which also may account for the discrepant findings of their study and Experiment 2. In the study by Koelsch et al. (2004), where longer excerpts of musical pieces (on average 10.5 s) were used in conjunction with a memory task, an N400 effect was found. It could therefore be hypothesized that, in the absence of a more elaborate processing of sound stimuli, the stimuli would need to be longer in order to elicit meaningful representations and thus an effect in the N400, whereas a longer sound stimulus would probably be sufficient to elicit meaningful associations even when elaborate processing is not required. The observed discrepancies between Experiment 1 and Experiment 2 could consequently be explained by two interacting factors—time of exposure and depth of processing. It seems plausible that a meaningful context may be established through either elaborate processing—by varying the task instructions—or through a longer presentation time or both. These issues need to be addressed in future studies. The absence of an N400 effect in Experiment 2 suggests that, at least with short sounds, meaningful representations are not automatically activated. The present study thus leaves open the question of whether meaningful representations are automatically activated in every type of musical listening experience.

Another factor potentially underlying the different outcomes of Experiment 1 and Experiment 2 (i.e., no N400 effect with a memory task) might be strategies used by the participants to solve the task. It is possible that participants tried to (sub-vocally) produce matching words for the sound stimuli, and thus effectively performed a semantic judgment for words only. This strategy could have been applied to all types of stimulus combinations, even sound-sound pairs, but would very likely not have been used for the memory task in Experiment 2, since participants did not have to perform a judgment on the relatedness of the stimulus pairs.

One drawback of the present study was the fact that the sounds differed in length and are also not perceived at a specific point in time (e.g., because some sounds unfold over time, whereas others do not). That is, in contrast to words, the sounds used in the present study varied in how they evolved and changed over time because they were characterized by different attack times, sustained portions, and decays. Therefore, it seems likely that these characteristics influence the time course of accessing the meaningful representation connected with a sound, and it is possible that activation of a meaningful representation connected to a sound with a short attack is quicker than to one with a longer attack, even if both sounds have the same length. This could potentially result in a temporally jittered N400 effect because the onset of the N400 might differ considerably between sounds. However, an effect in the N400 was still found for sound targets, showing that the N400 effect is nevertheless stable enough to be observed in the ERPs.

One known pathway to the expression of extra-musical meaning in music is emotion (Hevner, 1936; Sollberger et al., 2003; Steinbeis & Koelsch, 2008). The experiments presented here were aimed at exploring an alternative and independent pathway to meaning in music. The timbre of these sounds was found to influence the meaningful representations we have of them. This relationship between low level processing of musical features and meaning in music fits into Koelsch and Siebel's (2005) model of music perception. In this model, meaning in music cannot only arise when a musical piece is fully "understood"; that is, once all its structural aspects have been analyzed by the listener. Rather, at any point during a number of intermediate processing steps, meaningful processing of the musical information may take place. Basic level features of musical sounds, such as pitch height, pitch chroma, loudness, roughness, and timbre are extracted in early processing stages. All of these low-level features of music are thought to have a direct link to meaningful representations. This model is thus supported by the present observations that even short sounds, presented outside of a musical context, are able to activate representations of meaningful concepts. In this regard, Experiment 1 provides some evidence for a pathway to the access of extra-musical meaning. Basic level features of music, i.e., single sounds with a characteristic timbre, can access meaningful representations, even when presented outside of a musical context. However, the exact role of these basic level features when they occur in a musical context and how they interact with other musical features remain to be determined.

In conclusion, the present study shows that single sounds can activate representations of meaningful concepts in a similar fashion to chords and musical excerpts. No musical context is necessary to activate these representations. However, the task was found to have a great influence on the presence of an N400 effect. The activation of meaningful representations seems to occur via a distinct pathway, independently of the emotional valence of the sounds. Moreover, the distribution and time window of the N400 effect found in this experiment is comparable to that in other domains such as environmental sounds or language.

REFERENCES

Anderson, J., & Holcomb, P. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, *32*, 177–190.

Bentin, S., Kutas, M., & Hillyard, S. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology*, *30*, 161–169.

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

Chwilla, D., Brown, C., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, *32*, 274–285.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.

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

Deacon, D., Breton, F., Ritter, W., & Vaughan, H. (1991). The relationship between N2 and N400: Scalp distribution, stimulus probability, and task relevance. *Psychophysiology*, *28*, 185–200.

Daltrozzo, J., & Schoen, D. (2009a). Conceptual processing in music as revealed by N400. *Jounal of Cognitive Neuroscience*, *21*, 1882–1892.

Daltrozzo, J., & Schoen, D. (2009b). Is conceptual processing in music automatic? An electrophysiological approach. *Brain Research*, *1270*, 88–94.

Grigor, J., VanTolle, S., Behan, J., & Richardson, A. (1999). The effect of odor priming on long latency visual evoked potentials of matching and mismatching objects. *Chemical Senses*, *24*, 137–144.

Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERP to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, *113*, 1399–1350.

Heil, M., Rolke, B., & Pecchinenda, A. (2004). Automatic semantic activation is no myth: Semantic context effects on the N400 in the letter-search task in the absence of response time effects. *Psychological Science*, *15*, 852–857.

Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, *48*, 246–268.

Huynh, H., & Feldt, L. S. (1976). Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics*, *1*, 69–82.

Hohlfeld, A., & Sommer, W. (2005). Semantic processing of unattended meaning is modulated by additional task load: Evidence from electrophysiology. *Cognitive Brain Research*, *24*, 500–512.

Jasper, H. (1958). Report of the committee on methods of clinical examination in electroencephalography 1957. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375.

Kendall, R., & Carterette, E. (1993a). Verbal attributes of simultaneous wind instrument timbres: I. Von Bismarck's adjectives. *Music Perception*, *10*, 445–468.

Kendall, R., & Carterette, J. (1993b). Verbal attributes of simultaneous wind instrument timbres. II. Adjectives induced from Piston's Orchestration. *Music Perception*, *10*, 469–502.

Koelsch, S., & Siebel, W. (2005). Towards a neural basis of music perception. *Trends in Cognitive Science*, *9*, 578–584.

Koelsch, S., Gunter, T., Wittforth, M., & Sammler, D. (2005). Interaction between syntax processing in language and music. *Journal of Cognitive Neuroscience*, *17*, 1565–1577.

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.

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, *4*, 463–470.

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

Kutas, M., & Hillyard, S. (1989). An electrophysiological probe of incidental semantic association. *Journal of Cognitive Neuroscience*, *1*, 38–49.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology*, *36*, 53–65.

Meyer, L. (1956). *Emotion and meaning in music*. Chicago, IL: University of Chicago Press.

Nussbaum, C. (2007). *The musical representation. Meaning, ontology and emotion*. Cambridge, MA: MIT Press.

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

Patel, A. (2008). *Music, language and the brain*. New York, NY: Oxford University Press.

Raffman, D. (1992). Proposal for a musical semantics. In M. Riess Jones & S. Holleran (Eds.), *Cognitive bases for musical communication* (pp. 23–45). Washington, DC: APA Press.

Raffman, D. (1993). *Language, music and mind*. Cambridge, MA: MIT Press.

Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception*, *3*, 263–282.

Solomon, L. (1958). Semantic approach to the perception of complex sounds. *The Journal of the Acoustical Society of America*, *30*, 421–425.

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*. doi: 10.1371/journal.pone.0002226

Steinbeis, N., & Koelsch, S. (in press). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience*. doi: 10.1162/jocn.2009.21383

Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved February 20, 2010 from http://work-learning.com/effect_sizes.htm

van Petten, C., & Luka, B. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, *97*, 279–293.

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

(Received August 31, 2009; Accepted July 8, 2010)