

Perceptual and Conceptual Priming of Environmental Sounds

Annett Schirmer, Yong Hao Soh, Trevor B. Penney, and Lonce Wyse

Abstract

■ It is still unknown whether sonic environments influence the processing of individual sounds in a similar way as discourse or sentence context influences the processing of individual words. One obstacle to answering this question has been the failure to dissociate perceptual (i.e., how similar are sonic environment and target sound?) and conceptual (i.e., how related are sonic environment and target?) priming effects. In this study, we dissociate these effects by creating prime–target pairs with a purely perceptual or both a perceptual and conceptual relationship. Perceptual prime–target pairs were derived from perceptual–conceptual pairs (i.e., meaningful environmental sounds) by shuffling the spectral composition of primes and targets so as

to preserve their perceptual relationship while making them unrecognizable. Hearing both original and shuffled targets elicited a more positive N1/P2 complex in the ERP when targets were related to a preceding prime as compared with unrelated. Only related original targets reduced the N400 amplitude. Related shuffled targets tended to decrease the amplitude of a late temporo-parietal positivity. Taken together, these effects indicate that sonic environments influence first the perceptual and then the conceptual processing of individual sounds. Moreover, the influence on conceptual processing is comparable to the influence linguistic context has on the processing of individual words. ■

INTRODUCTION

Do we anticipate future auditory events based on what we hear in our sonic environment? To answer this question, imagine the sound of a door handle turning. If you then try to imagine a subsequent sound, it would likely be the opening of a door—the most probable outcome of operating a door handle. Whereas examples such as this illustrate our ability to infer event probabilities from sounds, they also raise the question of how the human brain generates such inferences. The answer to this question is not fully known. What we do know is that the processing of environmental sounds shares basic aspects with the processing of music and speech. All rely on a common neuronal pathway spanning from the ear to the primary auditory cortex. However, whether and how processing beyond this pathway differentiates between the sound classes are yet to be fully understood.

Evidence for a Differentiation of Sound Processing

Some researchers have used surface differences between environmental and other sounds to advocate for higher-order processing differences. For example, Van Petten and Rheinfielder (1995) highlighted the abstract nature of language and the fact that words map arbitrarily to meaning. In contrast, the mapping between environmental

sounds and meaning is not arbitrary but results from the physical properties of the eliciting event. Additionally, words and musical melodies are less source specific than environmental sounds. For example, words and melodies can be recognized independently of whether they are produced by a male or female speaker despite stark acoustic variation. In contrast, environmental sounds bear a more intimate relationship to their source (e.g., glass vs. plastic), such that changing the source generally means changing the sound. At the same time, environmental sounds differ in causal uncertainty from other sounds. There are many more similar sounds that have different causes than there are similar sounding words with different meanings. For example, many events produce clicking sounds (e.g., closing a button, operating a ball-point pen, and opening a lock), and it is, thus, difficult and perhaps impossible to infer the correct event from such a sound. In contrast, the same inference is less problematic from a verbal phrase describing the event (Ballas, 1993).

Differences in the processing of environmental and other sounds have also been inferred from neuroimaging and neuropsychological studies, which imply partially dissociated brain regions in the higher-order analysis of sounds. In an fMRI study, Noppeney, Josephs, Hocking, Price, and Friston (2008) observed that, although both the contextual processing of environmental sounds and speech evoked activity in the left inferior and medial superior frontal gyrus, activity in the left angular gyrus was specific to environmental sounds whereas activity in the left superior and

middle temporal gyrus was specific to speech. Slightly different processing differences were reported in simple word or sentence and sound processing studies that did not manipulate the stimuli's contextual relationship. Here, researchers observed left anterior and posterior superior temporal activations to speech relative to environmental sounds and right frontal or temporal activations for environmental sounds relative to speech (Thierry, Giraud, & Price, 2003; Humphries, Willard, Buchsbaum, & Hickok, 2001). A comparable pattern of lateralization was observed in research using ERPs (Van Petten & Rheinfelder, 1995). Finally, a case study of auditory agnosia provides evidence for a dissociation of environmental sound and speech processing. Specifically, Saygin, Leech, and Dick (2010) describe a patient with infarct damage to left temporal and parietal regions. After recovering from the infarct, the patient also recovered his speech, such that no impairments could be observed at the time of testing. However, his recognition of environmental sounds remained severely impaired. Although Saygin and colleagues advanced explanations other than a neuronal dissociation between speech and sound processing, at first glance, this and the preceding evidence imply that speech processing relies on especially dedicated mechanisms—mechanisms that are not accessed by environmental sounds.

Evidence for Commonalities in Sound Processing

Opposing this evidence is a large literature outlining comparable processing mechanisms and representations of the different sound classes. The core of this literature comprises priming studies testing whether the conceptual processing of environmental sounds compares with that of speech and music. In these studies, environmental sounds are typically presented with a word or a picture that portrays the source of the sound or that portrays an unrelated source. For example, the sound of a dog barking may be presented together with a verbal description or an image of a dog in the related condition and that of a car in the unrelated condition. If the time interval between both stimuli exceeds several seconds and is filled with other events, no advantage for the processing of congruous over incongruous pairs is observed (Friedman, Cycowicz, & Dziobek, 2003; Stuart & Jones, 1995). However, if the related source stimulus follows in immediate succession, researchers consistently observe priming. Specifically, behavioral decisions to related stimuli are faster and more accurate than behavioral decisions to unrelated stimuli. This is true both when environmental sounds serve as primes and behavioral decisions are measured to a word or picture as the target (Chen & Spence, 2010; Schneider, Engel, & Debener, 2008; Orgs, Lange, Dombrowski, & Heil, 2006; Van Petten & Rheinfelder, 1995) and when words or pictures serve as primes and behavioral decisions are measured to the environmental sound as the target (Schneider et al., 2008; Oezcan & van Egmond, 2003; Stuart & Jones, 1995; Ballas et al., 1993).

Apart from behavioral decisions, researchers have also measured the neuronal signatures of conceptual sound processing during cross-modal priming using ERPs. These experiments modulated the N400, an ERP component typically observed in studies of language processing and taken as an index of the retrieval of semantic information from memory. Words presented in the context of semantically related words were found to elicit a smaller N400 than words presented in the context of semantically unrelated words (Kutas & Hillyard, 1980). More recently, similar effects have been obtained for words following related and unrelated sounds (Schön, Ystad, Kronland-Martinet, & Besson, 2010; Daltrozzo & Schön, 2009; Koelsch et al., 2004; Van Petten & Rheinfelder, 1995) and for sounds following related and unrelated words or pictures (Schön et al., 2010; Daltrozzo & Schön, 2009; Orgs, Lange, Dombrowski, & Heil, 2008; Cummings et al., 2006; Orgs et al., 2006; Plante, Van Petten, & Senkfor, 2000; Van Petten & Rheinfelder, 1995). On the basis of this evidence, researchers proposed that, like speech, environmental sounds can be processed conceptually. Moreover, they argued that the way conceptual information is stored and used is comparable across the different sound classes.

However, a closer look at the available evidence suggests these conclusions to be premature. To date, the studies demonstrating an N400 effect for the processing of sounds other than speech all employed cross-modal paradigms. A word or a picture served either as the prime or the target. This mode of testing undoubtedly induced participants to generate a common code for stimulus evaluation. Although it is possible that this code was an amodal conceptual representation, it is equally possible that it was verbal both when environmental sounds were presented with words and when they were presented with pictures. In the case of words, only the sounds would have to be verbalized for easy matching. In the case of pictures, a common verbal code may be more readily available than an auditory code for a picture or a visual code for a sound. If this were the case, then this verbal code rather than the natural conceptual processing of environmental sounds could account for the observed N400 effect.

That this is a valid concern can be easily appreciated when considering findings from another sound class. Specifically, in the case of music, researchers observed a cross-modal N400 effect, similar to the one reported for environmental sounds (Daltrozzo & Schön, 2009; Koelsch et al., 2004). Short musical excerpts (e.g., from marching music) were found to reduce the N400 of subsequently presented words if they were conceptually related (e.g., courage) as compared with unrelated (e.g., magic). Furthermore, words have been shown to reduce the N400 of conceptually related as compared with unrelated musical excerpts. However, to our knowledge, no published study has shown such an effect within music alone. Studies that introduced expectancy violations within music revealed different ERP effects. For example, distantly related chords have been found to elicit an early negativity (Schön, Magne,

& Besson, 2004; Koelsch, Gunter, Friederici, & Schröger, 2000) and unexpected notes elicit a late positive shift (Miranda & Ullman, 2007; Schön et al., 2004). One may argue that these violations were qualitatively different from the conceptual violations introduced in the cross-modal priming experiments. This, however, raises two questions. Why was conceptual processing not affected by these structural violations and what prevents music researchers from creating within-modality conceptual manipulations that elicit an N400 effect? A demonstration of such an effect appears vital for establishing conceptual processing similarities between language, music, and environmental sounds.

The Present Study

Taken together, the evidence on whether the higher-order analysis of environmental sounds compares with that of other sounds such as speech and music is divided. On the one hand, there is evidence for differences in the underlying neuronal substrates. On the other hand, behavioral and ERP priming studies raise the possibility of comparable processing mechanisms. We tested the latter possibility using a within-modality priming paradigm in conjunction with sound pairs that were conceptually related or unrelated. When creating these sound pairs, we considered the possibility that any conceptual relationship could be confounded by a perceptual relationship. For example, the typing of a typewriter is conceptually related to the “ding” heard at a carriage return. However, it is possible for primes and targets to share perceptual properties and, thus, sound similar. In this case, any processing differences between related and unrelated pairs could be conceptual, perceptual, or both. To disentangle this, we created control sound pairs. These were derived from the original sound pairs by shuffling the phase of their spectral components. This produced sounds that largely preserved the perceptual relationship between primes and targets while rendering them unrecognizable. Pairs of original and shuffled sounds were presented to participants with the instruction to listen carefully as sound memory would be tested subsequently.

We predicted that the relatedness between prime and target sounds would influence the ERP to targets. Perceptual relatedness was expected to influence ERP components previously associated with acoustic or probabilistic processing. As such, we anticipated a modulation of the N1/P2 complex and a late positive shift, both of which have been previously observed in studies manipulating sound expectancy (Miranda & Ullman, 2007; Schön et al., 2004; Koelsch et al., 2000). Given that the perceptual relationship between primes and targets was comparable for the original and shuffled versions, the expected ERP modulations should show in either case, thus demonstrating their independence from sound recognizability. Conceptual relatedness, in contrast, should depend on sound recognizability and modulate the processing of original targets only. Moreover, if sound processing evokes language-like conceptual

mechanisms, original but not shuffled targets should produce a larger N400 when they are unrelated as compared with related.

EXPERIMENT 1

Methods

Participants

Eighteen participants took part in the experiment. The data from two participants were excluded from the analysis because of excessive eye-blink artifacts. Eight of the remaining participants were men, and eight were women. Ages ranged from 19 to 24 years ($M = 21.8$), and all participants reported normal hearing and normal or corrected-to-normal vision. All participants provided written informed consent before the experiment.

Stimuli

We created 52 sound sequences that were deemed conceptually related and 52 sequences that were deemed unrelated. The primes in both sequences were identical, and the targets differed. The primes were relatively longer than the targets, such that they would create a target expectation similar to sentence context. The sounds were subjected to two rating studies. The first study comprised an individual sound rating. Here, 30 participants, who did not contribute to the main study, listened to each sound played in isolation. Following the presentation of each sound, the participant indicated whether she or he could recognize it. If the answer was positive, the participant wrote down its source and event. For example, when recognizing the typing of a typewriter, the participant had to write down “typewriter” and “typing.” After submitting a positive or negative recognition answer, the participant saw four source options and the option “none of the above” and had to indicate the correct source. Unbeknown to the participant, “none of the above” was never the correct answer but provided a response option in case the participant failed to recognize the sound. This screen was followed by four event options and the option “none of the above,” and the participant now had to identify the correct event. Finally, the participant was asked to rate the familiarity of a given sound on a scale from 1 (*unfamiliar*) to 5 (*highly familiar*).

The second rating study assessed sounds in pairs. Again 30 individuals who did not participate in the main experiment were recruited. Each listened to the prime of a given sound pair and the target that followed the prime by 500 msec and subsequently rated the perceptual similarity of prime and target on a scale from 1 (*dissimilar*) to 5 (*very similar*). Following this rating, the participant assessed the conceptual relationship between prime and target on a scale from 1 (*unrelated*) to 5 (*highly related*).

The rating results were used to select a subset of prime–target pairs for the EEG study. This subset comprised 27 primes with two sets of targets: 27 conceptually related

and 27 conceptually unrelated. The rating results are presented in Tables 1 and 2. The selected prime sounds were, on average, 4500 msec ($SD = 2782$) in duration. The selected target sounds were all 800 msec in duration.

We introduced a sound memory task to keep participants' attention focused on the sounds during the experiment proper (for more details, please refer to Paradigm). Thus, in addition to the experimental sounds, we selected 54 sounds to serve as new sounds in a sound recognition test performed after the main experiment. The 54 sounds were comparable to the targets in duration, intensity, recognizability, and familiarity.

As illustrated in Table 2, both the perceptual and the conceptual rating scores were significantly greater for related as compared with unrelated targets. This could mean that participants failed to discriminate the two constructs. Alternatively, it could imply that conceptually related sounds tend to also be perceptually similar. This is possible because they might be generated under similar acoustic conditions (e.g., same room) or from the same source (e.g., typewriter) and, as a consequence, share acoustic information (e.g., frequency content or temporal envelope). As the primary focus of this study was to investigate conceptual sound processing, we were concerned about a potential perceptual confound. To address this confound, we created control sound pairs that matched the original sound pairs in perceptual relatedness but were unrecognizable, hence should not induce conceptual priming. Similarly, we created control sounds for the 54 new sounds used in the sound recognition test.

This was achieved as follows. First, we determined the temporal envelope for each original prime and target sound as well as the 54 new sounds. Then, all these sounds were subjected to a Fourier transform, which decomposed the sound signal into individual weighted sine waves (e.g., Tae, 2010). Then, the phase information of each sine wave was shuffled, and the inverse Fourier transform was used to generate a complex wave from the sine waves with the shuffled phases and their original weights. As the phase shuffling distorted the original temporal envelope of a given sound, the obtained complex

Table 1. Sound Recognition

Sound Type	Source				Event			
	Written		Multiple Choice		Written		Multiple Choice	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Prime	1.62	0.36	0.95	0.07	1.69	0.31	0.88	0.15
Related	1.29	0.52	0.89	0.14	1.26	0.47	0.86	0.11
Unrelated	1.26	0.53	0.91	0.12	1.32	0.45	0.88	0.14
New Sounds	1.30	0.52	0.92	0.08	1.26	0.42	0.88	0.09

The maximum score for the written task is 2. The maximum score for the choice task is 1.

Table 2. Relatedness Ratings for Original Prime–Target Pairs

Sound Type	Conceptual		Perceptual	
	Mean	SD	Mean	SD
Related	4.21	0.38	2.81	0.57
Unrelated	1.45	0.52	1.99	0.71

The score for maximal relatedness is 5.

wave was multiplied with the temporal envelope of the original sound (for similar algorithm, cf. Kirmse, Jacobsen, & Schröger, 2009; Altmann, Doehrmann, & Kaiser, 2007; Lenz, Schadow, Thaerig, Busch, & Herrmann, 2007). This produced sounds that were unrecognizable but shared frequency content and envelope with the original sounds. Hence, the perceptual relationship between primes and targets, as reflected in frequency content and envelope, was comparable for original and shuffled stimuli. Exemplary spectrograms of original and shuffled prime–target pairs are presented in Figure 1.

Paradigm

The participant sat in a comfortable chair facing a monitor at a distance of about 60 cm. The experiment comprised a study phase, during which the participant listened to sound pairs over head phones, and a test phase, in which the target sounds and a set of new sounds were presented and the participant indicated whether a given sound was old or new.

In the study phase, a trial started with a white fixation cross presented in the center of a black screen. After 1000 msec, the prime appeared and, after a 500-msec ISI, was followed by the target. The fixation cross disappeared with target offset, and the next trial began after either 500, 1000, or 1500 msec, with one third of the trials occurring at each ITI duration. Participants were asked to refrain from blinking during the presentation of the cross and to remember the sounds for a subsequent memory test. The study phase comprised 27 original/related prime–target pairs, 27 original/unrelated prime–target pairs, 27 shuffled/related prime–target pairs, and 27 shuffled/unrelated prime–target pairs. Each prime occurred twice, once followed by the related target and once followed by the unrelated target. The presentation of prime–target pairs was pseudo-randomized, such that no more than three trials of the same condition occurred in succession. Moreover, prime–target pairs were divided into two lists, such that a given prime occurred only once in each list. Both lists were presented with a short break in between to ensure that prime repetition was maximally spaced.

During the test phase, a trial again started with a white fixation cross presented in the center of the computer screen. After 1000 msec, a sound was presented. Following sound offset, the words “old” and “new” appeared on the

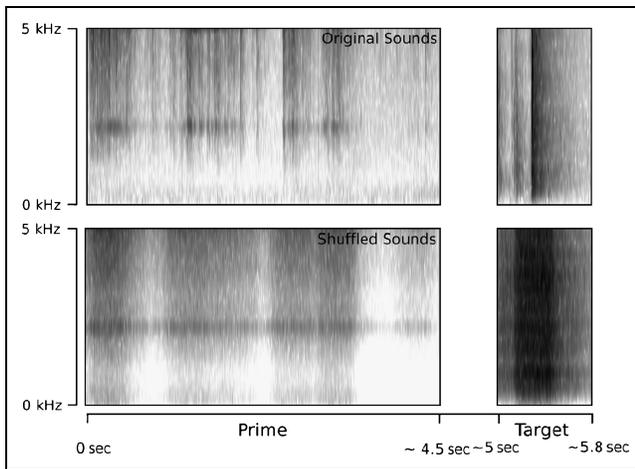


Figure 1. Example spectrograms of original prime (i.e., jingling keys) and target (i.e., door unlocking) are shown in the first row. Their phase shuffled versions are shown in the second row.

left and right side of the screen, and the participant was required to press the corresponding left or right button on the response box indicating whether a sound was heard during the study phase. The next trial started after 500, 1000, or 1500 msec, with one third of the trials occurring at each ITI duration. The test phase comprised 216 trials, half of which used targets from the preceding study phase and half of which used original and shuffled new sounds.

Electrophysiological Recording and Analysis

The EEG was recorded from 64 electrodes mounted in an elastic cap according to the modified 10–20 system. The EOG was recorded from three electrodes attached above and below the right eye and at the outer canthus of the left eye. Additionally, one recording electrode was placed on the nose tip and one on each mastoid. The data were recorded at 256 Hz with an ActiveTwo system from Biosemi (Amsterdam, the Netherlands), which uses a common mode sense active electrode for initial referencing.

EEG/EOG data were processed with EEGLAB (Delorme & Makeig, 2004). The recordings were re-referenced to the nose and a 0.5- to 20-Hz bandpass filter was applied. The continuous data were epoched and baseline-corrected using a 200-msec prestimulus baseline and a 1000-msec time window starting from stimulus onset. Nontypical artifactual epochs caused by drifts or muscle movements were rejected automatically. Infomax, an independent component analysis algorithm implemented in EEGLAB, was applied to the remaining data, and components reflecting typical artifacts (i.e., horizontal and vertical eye movements and eye blinks) were removed. Back-projected single trials were screened visually for residual artifacts.

Only the ERPs to targets were analyzed. These were derived by averaging individual epochs for each condition and participant. Of interest for the present study were the

N1, P2, N400, and the late positive component (LPC). We identified the N1 and N400 peak by determining the most negative value between 100 and 200 msec and between 300 and 500 msec, respectively. We identified the P2 and LPC peak by determining the most positive value between 150 and 250 msec and between 500 and 700 msec, respectively. A 50-msec (for N1 and P2) or a 200-msec (for N400 and LPC) time window was centered around component peaks, and mean values in these time windows were subjected to separate ANOVAs with *Sound* (Original/Shuffled), *Relatedness* (Related/Unrelated), *Region* (Anterior/Posterior), and *Laterality* (Left, Center, Right) as repeated measures factors. The factors *Region* and *Laterality* comprised the following subgroups of electrodes: anterior left, AF3 AF7 F3 F7; anterior middle, FPZ AFZ FZ FCZ; anterior right, AF4 AF8 F4 F8; posterior left, P3 P7 PO3 PO7; posterior middle, CPZ PZ POZ OZ; posterior right, P4 P8 PO4 PO8. This selection of electrodes ensured that the tested subgroups contained equal numbers of electrodes while providing a broad scalp coverage that enabled assessment of topographical effects.

Perceptual priming was expected to reveal a *Relatedness* main effect that was unqualified by *Sound*. This is because such an effect would be sensitive to what was common between the original and shuffled prime–target pairs—that is their perceptual relation. Conceptual priming was expected to reveal an interaction of *Relatedness* and *Sound* as it should show for original prime–target pairs only. Where appropriate, *p* values were corrected for sphericity using the Greenhouse–Geisser procedure (p_{GG}).

Results

Behavioral Results

The memory task served primarily as a means to engage participants in active sound processing, and the results from the task are illustrated in Figure 2. Overall recognition

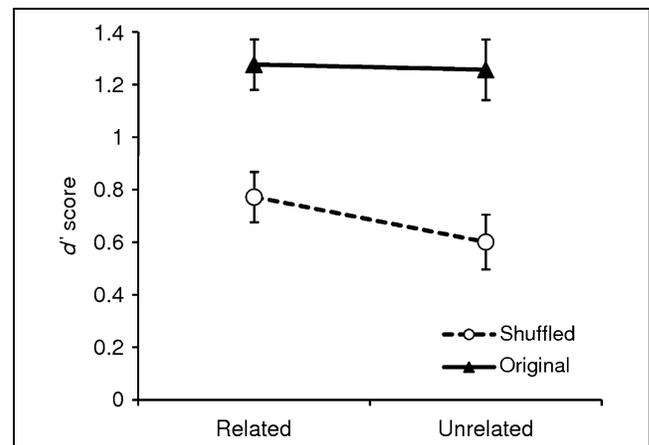


Figure 2. Memory scores for original and shuffled sounds in Experiment 1.

performance was assessed by computing the d' discrimination index for each participant. An ANOVA treating *Sound* and *Relatedness* as repeated measures factors revealed a *Sound* main effect ($F(1, 15) = 33.1, p < .0001$), which indicates better memory for original as compared with shuffled sounds. All other effects were nonsignificant (all $ps > .1$).

EEG Results

The ERP results are illustrated in Figures 3 and 4. The N1 peaked 150 msec following stimulus onset ($SD = 28$). Analysis of mean voltages between 125 and 175 msec following stimulus onset revealed marginally significant effects of *Sound* ($F(1, 15) = 3.4, p = .08$) and *Relatedness* ($F(1, 15) = 4.1, p = .06$) with larger N1 amplitudes to original relative to shuffled sounds and unrelated relative to related sounds.

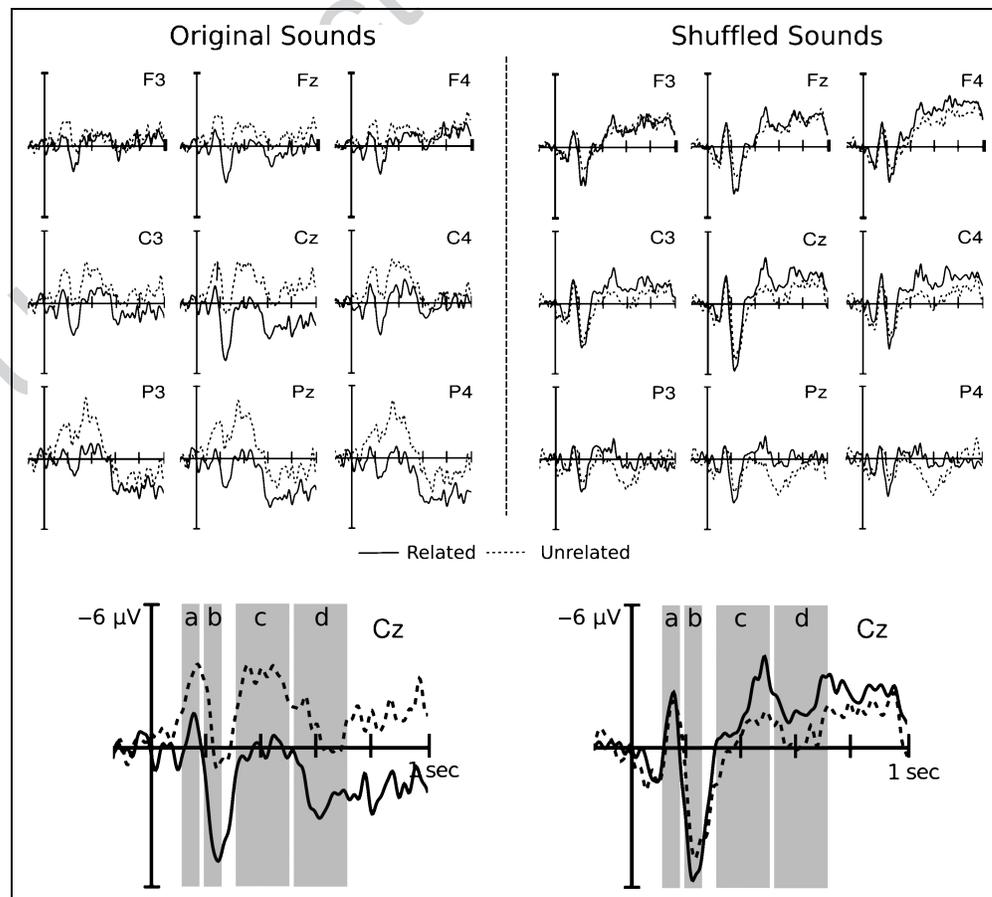
The P2 peaked 209 msec following stimulus onset ($SD = 31$). Analysis of mean voltages between 184 and 234 msec following stimulus onset revealed a significant main effect of *Sound* ($F(1, 15) = 24.1, p < .001$), indicating larger P2 amplitudes for shuffled sounds relative to original sounds. Additionally, there was a *Relatedness* main effect ($F(1, 15) = 12.9, p < .01$), indicating larger P2 amplitudes for related as compared with unrelated sounds. This latter

effect was further qualified by an interaction with *Laterality* ($F(2, 30) = 13.1, p_{GG} < .0001$). Follow-up analysis of the interaction revealed that the *Relatedness* effect was larger over midline ($F(1, 15) = 20.2, p < .001$) as compared with left ($F(1, 15) = 5, p < .05$) and right ($F(1, 15) = 15.1, p < .01$) hemisphere regions.

The N400 peaked 398 msec following stimulus onset ($SD = 66$). Analysis of mean voltages between 298 and 498 msec following stimulus onset revealed an interaction of *Sound*, *Relatedness*, and *Region* ($F(1, 15) = 6.3, p < .05$) and an interaction of *Sound*, *Relatedness*, and *Laterality* ($F(1, 15) = 4.3, p_{GG} < .05$). Follow-up analysis for original sounds revealed a *Relatedness* by *Region* ($F(1, 15) = 4.3, p < .05$) and a *Relatedness* by *Laterality* interaction ($F(2, 30) = 8.5, p_{GG} < .01$). Over right ($F(1, 15) = 4.6, p < .05$), middle ($F(1, 15) = 5.8, p < .05$), and posterior sites ($F(1, 15) = 5.7, p < .05$), unrelated sounds elicited a greater N400 than related sounds. No other effects reached significance (all $ps > .1$). The follow-up analysis for shuffled sounds was nonsignificant (all $ps > .1$).

The LPC peaked 602 msec following stimulus onset ($SD = 58$). Analysis of mean voltages between 502 and 702 msec following stimulus onset revealed a significant interaction between *Sound*, *Relatedness*, and *Region* ($F(1, 15) = 6.6, p < .05$). Follow-up analysis for original sounds was nonsignificant ($p > .1$). Follow-up analysis for shuffled sounds

Figure 3. Grand average ERPs for original and shuffled sounds in Experiment 1. The waveforms are time-locked to target onset.



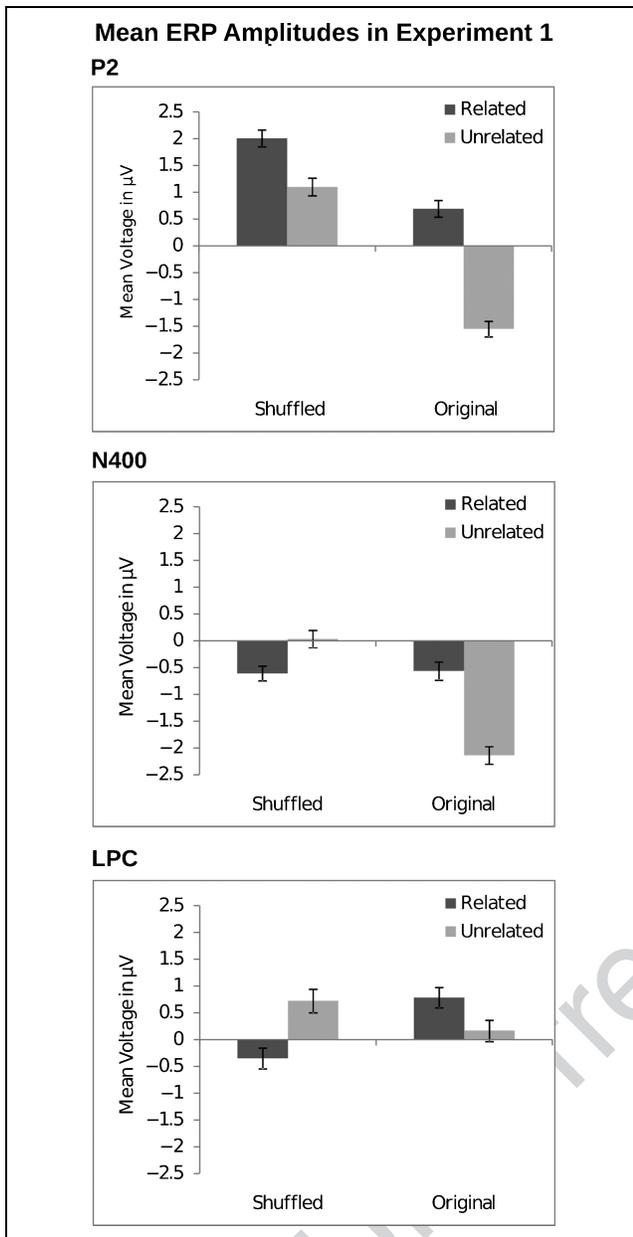


Figure 4. Mean ERP amplitudes in time windows of statistical analysis. Error bars reflect SEM.

revealed a *Relatedness* by *Region* interaction ($F(1, 15) = 4.5, p < .05$), indicating that ERPs tended to be more positive in the unrelated as compared with the related condition over posterior sites ($F(1, 15) = 4.2, p = .058$). All other effects were nonsignificant ($p > .1$).

EXPERIMENT 2

Experiment 1 revealed a number of effects indicative of the perceptual and conceptual relationship between sound primes and targets. Related sounds elicited a marginally smaller N1 and a larger P2 relative to unrelated sounds, regardless of whether they were original or shuffled. Original sounds additionally showed an N400 relatedness effect,

whereas shuffled sounds showed a marginally significant LPC relatedness effect. However, because the targets in the related and unrelated conditions were, in fact, different sounds, one may object that any ERP differences between them reflect basic sound differences rather than priming effects. To address this concern, we conducted a second experiment in which the targets were presented without the primes. Hence, any effects related to priming observed in Experiment 1 should no longer be present in this second experiment.

Methods

Participants

Eight male and eight female participants ranging in age from 20 to 24 years ($M = 21.8$) took part in the experiment. All participants reported normal hearing and normal or corrected-to-normal vision.

Stimuli, Procedure, Electrophysiological Recording, and Data Analysis

The stimuli, procedure, electrophysiological recordings, and data analysis were comparable to Experiment 1. The only difference was that target sounds were presented without primes.

Results

Behavioral Results

The d' discrimination index was computed for each participant, and mean values for each condition are illustrated in Figure 5. An ANOVA treating *Sound* (Original/Shuffled) and *Target Stimulus Set* (Exp1-Related, Exp1-Unrelated) as repeated measures factors revealed a significant main effect of *Sound* ($F(1, 15) = 80.4, p < .0001$), indicating better memory for original as compared with shuffled sounds. All other effects were nonsignificant (all $ps > .1$).

ERP Results

The ERP results are illustrated in Figure 6. The N1 peaked 142 msec following stimulus onset ($SD = 25$). As in the main experiment, a 50-msec time window was centered around the peak. Analysis of mean voltages within this time window revealed no significant effects (all $ps > .1$).

The P2 peaked 208 msec following stimulus onset ($SD = 26$). Analysis of mean voltages between 183 and 233 msec following stimulus onset revealed a significant main effect of *Sound*, indicating that the P2 was larger for shuffled relative to original sounds ($F(1, 15) = 6.5, p < .05$).

The N400 peaked 386 msec following stimulus onset ($SD = 68$). Analysis of mean voltages within a 200-msec time window centered around the N400 peak revealed a significant interaction of *Sound* and *Region* ($F(1, 15) = 7, p < .05$), indicating that the N400 was larger for original as

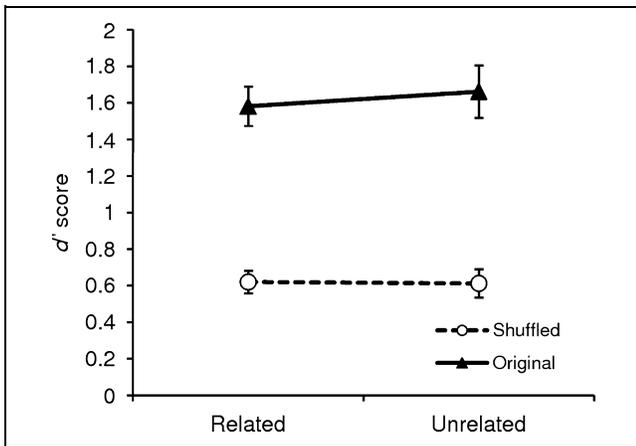
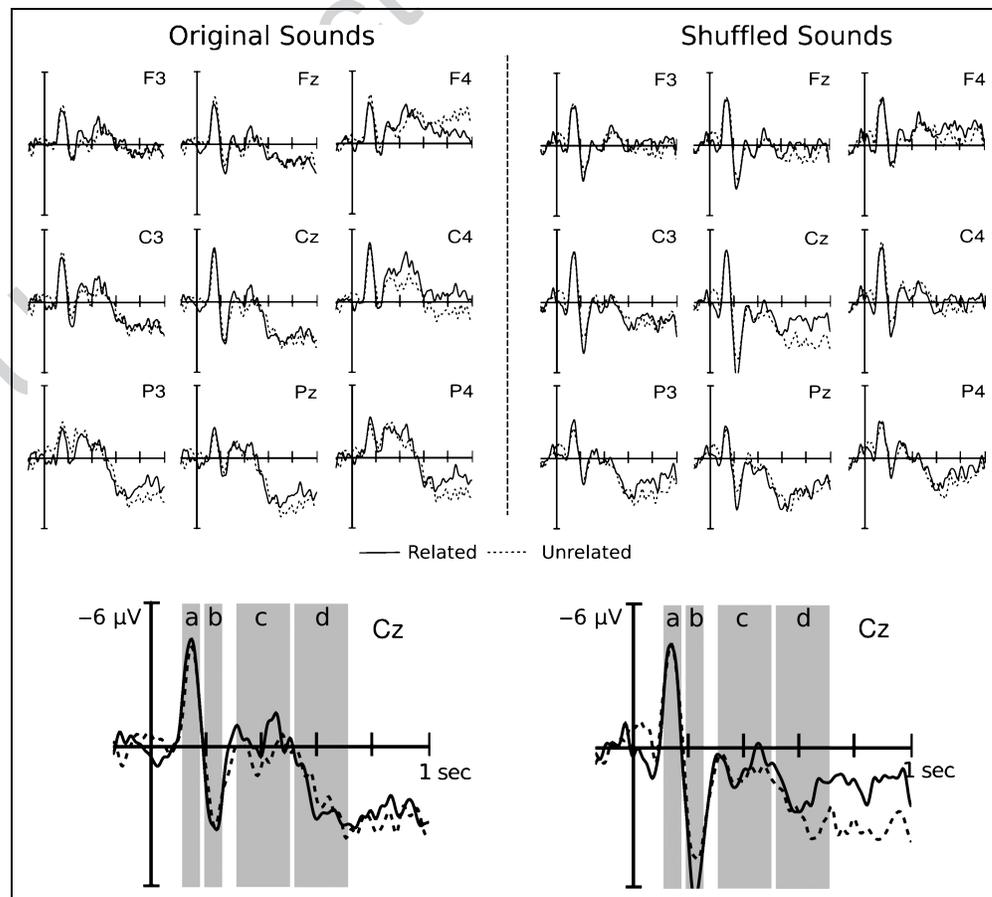


Figure 5. Memory scores for original and shuffled sounds in Experiment 2.

compared with shuffled sounds over posterior ($F(1, 15) = 7.1, p < .05$) but not anterior sites ($p > .1$). The interaction of *Sound*, *Target Stimulus Set*, *Hemisphere*, and *Region* was marginally significant ($F(2, 30) = 3.4, p_{GG} = .07$). Importantly, however, follow-up analyses for each region in the left, center, and right position indicated that the *Target Stimulus Set* main effect and the interaction between *Target Stimulus Set* and *Sound* were nonsignificant (all p s $> .5$). Moreover, the effect of the marginally

Figure 6. Grand average ERP for original and shuffled sounds in Experiment 2. The waveforms are time-locked to target onset.



significant interaction was in fact opposite to that observed in Experiment 1.

The LPC peaked 617 msec following stimulus onset ($SD = 56$). Analysis of mean voltages between 517 and 717 msec following stimulus onset revealed no significant effects (all p s $> .1$).

GENERAL DISCUSSION

The present study investigated whether and in what way the relationship between an environmental sound and its sonic context affects sound processing. In a rating study, we identified prime–target pairs that were conceptually—but also perceptually—related. These sounds, together with spectrally shuffled sounds, elicited a number of ERP effects. First, we observed effects associated with the nature of the sounds. Original sounds elicited more negative responses than shuffled sounds in the N1/P2 time range in Experiments 1 and 2 and in the N400 time range in Experiment 2. This suggests that original sounds recruited early sensory processing mechanisms to a greater extent than shuffled sounds (Schirmer et al., 2008; Sable, Low, Maclin, Fabiani, & Gratton, 2004) and were more likely to activate conceptual processing associated with the N400 (Kutas & Hillyard, 1980). Second and more importantly, we observed ERP effects associated with the relatedness of

targets to primes in Experiment 1. Given that these effects were absent in Experiment 2, when targets were presented without primes, they clearly reflect contextual processing. Moreover, given that some ERP effects were present for original *and* shuffled pairs whereas others were specific to one or the other, we can conclude that sonic context affects both perceptual and conceptual aspects of sound processing. These two aspects are discussed in more detail below.

Perceptual Sound Priming

Whether a sound primes other sounds perceptually has been investigated with repetition paradigms. In these paradigms, participants listened to a series of sounds. Some of the sounds had been heard previously, others had been heard as different exemplars, whereas others were new. For example, the ringing of a telephone could be repeated as the exact same ringing, the ringing of a different telephone, or a completely new sound could be presented. Generally, researchers observed faster and/or more accurate behavioral responses to an exact repetition of an item as compared with a new item. This comparison, however, entails both perceptual and conceptual repetition; thus, behavioral differences may be due to either aspect. To isolate perceptual facilitation, researchers have contrasted responses to exact repetitions and exemplar repetitions. In some cases, this failed to reveal differences (Stuart & Jones, 1995), whereas in other cases, greater benefits were observed for exact than exemplar repetition (Chiu, 2000).

The present study extends this work by elucidating the neurophysiological correlates and time course of perceptual priming. Specifically, we found related sounds elicited a more positive N1/P2 complex relative to unrelated sounds. Moreover, given that this effect occurred for both recognizable and nonrecognizable sounds (Figure 4), one may infer that the underlying mechanism was common to both, hence perceptual rather than conceptual in nature. This interpretation is in accord with the musical tone and chord expectancy findings described above (Miranda & Ullman, 2007; Schön et al., 2004; Koelsch et al., 2000) as well as findings from ERP investigations of stimulus repetition processing. For example, the repeated presentation of an auditory event in a sound sequence will result in successively smaller N1 and larger P2 amplitudes (Schirmer et al., 2008; Sable et al., 2004). This has been attributed to both the refractoriness of N1 neural generators and an inhibitory mechanism that reduces sensitivity to event repetition (Sable et al., 2004). Furthermore, in old or new recognition memory tests, previously encountered visual objects elicit a more positive ERP relative to new objects in the N1/P2 time range. This has been linked to perceptual facilitation and a sense of familiarity for test items that map onto a previously activated perceptual trace (Harris, Cutmore, O’Gorman, Finnigan, & Shum, 2009). The present results

imply similar mechanisms for the processing of perceptually related environmental sounds. Specifically, sounds that share acoustic properties with their sonic context engage some neural generators that were also engaged in context processing. Hence, these generators are refractory or inhibited. The resulting attenuation of auditory processing then likely creates a sense of familiarity or sound relatedness.

Apart from modulations in the N1/P2 range, the present study revealed a marginal effect for a late positive potential over temporo-parietal sites. This potential peaked between 500 and 700 msec, following target onset with larger amplitudes for unrelated as compared with related unrecognizable sounds. A similar effect has been reported for music (Schön et al., 2004) as well as other types of auditory sequences and is known as an “oddball” effect. Compared with standard or expected events, improbable and/or unexpected events in an attended event sequence elicit a larger late positive shift known as the P300. For example, an attended tone that deviates from preceding tones in pitch, intensity, or duration elicits a larger P300 than standard tones (Schröger, 1996). A visually processed word that differs from preceding words in font elicits a larger P300 relative to a word with comparable font (Kutas & Hillyard, 1980). As for the N1/P2 modulations, some have attributed these findings to an inhibitory mechanism. Presumably, this mechanism is activated when an event in the environment requires focused attention for processing (Polich, 2007). In this case, extraneous processes irrelevant for that event may be suppressed. The more suppression required, the larger the P300. Unlike the N1/P2 modulations, this late temporo-parietal effect is more controlled or reflective in nature, as it requires attention to be directed at the critical stimulus.

The late positive effect in the present study resembles a P300 and, thus, may be interpreted in a similar way. Unrelated sounds may have triggered a larger P300 than related sounds, because they were perceptually less familiar and required a stronger inhibition of extraneous processes. That the LPC/P300 difference was observed for unrecognizable, but not recognizable, sounds may be because it is more sensitive to perceptual than conceptual familiarity, and participants focused more strongly on perceptual features for the former relative to the latter. Such a difference in focus could arise because only perceptual features were available to support memory encoding of unrecognizable sounds. In contrast, perceptual and conceptual features could support memory encoding of recognizable sounds. As a consequence, perceptual violations in unrecognizable prime–target pairs may have been more salient and, thus, facilitated P300 generation. Alternatively, both unrecognizable and recognizable pairs elicited inhibition during both an early preattentive (i.e., N1/P2) and a late controlled processing stage (i.e., P300). However, for recognizable pairs, inhibition in the latter stage may have been disguised by a temporal overlap with conceptual processing in the N400 time range.

Conceptual Sound Priming

Some attempts to determine whether the conceptual processing of sounds matches that of speech simply compared the neuronal structures activated by sounds and speech. This work provided conflicting evidence with some studies suggesting a clear dissociation (Thierry et al., 2003; Humphries et al., 2001; Van Petten & Rheinfelder, 1995) and others implying a general overlap (Dick et al., 2007). Attempts to elucidate conceptual sound processing have also been made in cross-modal priming studies. Sounds have been used to prime pictures or words, and pictures or words have been used to prime sounds. Here, researchers consistently observed facilitated behavioral and neuronal responses when prime and target were conceptually related as compared with when they were unrelated (Schön et al., 2010; Daltrozzo & Schön, 2009; Orgs et al., 2006, 2008; Orgs, Lange, Dombrowski, & Heil, 2007; Cummings et al., 2006; Plante et al., 2000; Van Petten & Rheinfelder, 1995). However, although suggestive of conceptual sound processing, this evidence is not fully convincing.

First, it is possible that participants used a common verbal code to perform cross-modal matching. This concern was raised recently by Schön and colleagues (2010), who also investigated the conceptual processing of sounds in a cross-modal priming study. Unlike prior work, however, their study employed words together with natural sounds for which participants were unable to specify the source. Through this, the authors attempted to prevent verbalization of sounds and to force participants to access a stored concept without referring to language. Nevertheless, it is still possible that participants verbalized sound properties (e.g., harsh, metallic, and fast) in the effort to judge congruity between sounds and words and that this verbalization explains the observed conceptual priming effects.

A second issue is that, to date, studies using purely acoustic manipulations failed to convincingly demonstrate language-like conceptual processing (Chiu, 2000; Chiu & Schacter, 1995; Stuart & Jones, 1995; Ballas & Mullins, 1991). A few behavioral studies have reported processing facilitation using within-modality sound priming. Ballas and Mullins (1991) demonstrated that participants are better at identifying perceptually ambiguous sounds (e.g., ringing) when these sounds are presented in their proper sound context (e.g., telephone dialing). More recently, Schneider and colleagues (2008) used 400-msec-long sound clips in a series of cross- and within-modal priming experiments. In line with the evidence cited above, they found categorical decisions to be facilitated for related as compared with unrelated targets, regardless of whether only sounds or sounds and pictures were used. However, whereas the cross-modal manipulation raises coding issues, the within-modality manipulations failed to dissociate perceptual from conceptual priming. Thus, it is likely that conceptually related sound pairs also shared perceptual features that potentially facilitated behavioral decisions, a possibility that was addressed in the present study. Finally, research using

on-line measurements of neuronal activity to study the processing of unexpected or unpredicted auditory events in an auditory sequence has not yet revealed evidence for conceptual processing outside the realm of speech.

The present study provides such evidence and, for the first time, outlines the time course of perceptual and conceptual aspects of environmental sound processing. As discussed above, any relatedness effects observed for unrecognizable, shuffled sound pairs reflect perceptual processing, as these sounds have no shared conceptual representations. Such effects were observed between 100 and 200 msec and again around 500 msec following stimulus onset. More importantly, any relatedness effects observed for recognizable, original sounds that were absent for unrecognizable, shuffled sounds must be conceptual in nature. Such effects were observed between 300 and 500 msec following stimulus onset. In this time range, unrelated recognizable sounds elicited a greater N400 than related recognizable sounds over right-central and temporo-parietal regions. This is reminiscent of the well-documented N400 effect in language, which peaks at about the same time with a similar scalp topography. Notably, the right lateralization observed here matches that of the N400 obtained to visual word presentations (Holcomb & Neville, 1990). Although the present N400 distribution does not necessarily map onto the position of its underlying sources, it raises the possibility that the conceptual analysis of environmental sounds is less dependent on the fine temporal processes required for phoneme perception in speech that are known to be lateralized to the left hemisphere (for a review, see Schirmer, 2004). Moreover, both environmental sounds and visual words may recruit left hemisphere generators to a lesser extent resulting in a relative right lateralization of the ERP.

Although our study presents the first evidence that sound context, apart from speech, produces a within-modality N400 priming effect, it does not fully exclude the possibility that this effect resulted from verbalization. We may not have promoted verbalization by avoiding differences in code between prime and target. However, we also did not prevent verbalization. Nevertheless, that our findings are unlikely to reflect verbalization is suggested by a study of Dick, Bussiere, and Saygin (2002), who found participants to be slower in a sound matching task when asked to covertly name a sound as compared with not name it. Hence, it seems that verbalizing sounds adds processing effort that participants may not naturally engage in unless prompted to do so.

When considering the present evidence for conceptual sound processing, one may ask to what extent this evidence can be reconciled with the presumed neurofunctional specialization for language. As mentioned in the Introduction, some researchers have observed differences in the representation of linguistic and nonlinguistic stimuli. For example, differences have been reported in the activation of brain structures for word and sound processing (Noppeney et al., 2008; Van Petten & Rheinfelder, 1995). Moreover, there is evidence from a case study of auditory agnosia that

the brain dissociates speech from environmental sounds (Saygin et al., 2010). At this point, it is unclear how to interpret this evidence. Different functional neuroimaging patterns for spoken words and sounds may reflect methodological difficulties in matching the concepts represented by both. We know that concepts are organized both semantically and perceptually (Thompson-Schill, Aguirre, D'Esposito, & Farah, 1999; Farah, Hammond, Mehta, & Ratcliff, 1989). For example, concepts related to actions (e.g., typing) rely more strongly on motor representations than concepts related to objects (e.g., typewriter). Thus, the former are more likely to recruit premotor or motor areas as compared with the latter (Galati et al., 2008; Pizzamiglio et al., 2005). Past research failed to equate the semantic and perceptual attributes of word and sound stimuli. For example, the words used typically constituted nouns and, thus, referred to objects, whereas environmental sounds are by nature also indicative of an action or event. It is, thus, possible that the poor match between nouns and environmental sounds introduced processing differences that are not reflective of actual differences in the representation of speech and environmental sounds.

Furthermore, the auditory agnosia case cited earlier (Saygin et al., 2010) stands out from among many cases in which both the processing of speech and that of environmental sounds were affected (Saygin, Dick, Wilson, Dronkers, & Bates, 2003). Hence, those authors do not argue that the two are truly independent. Instead they suggest that, depending on the lesion, fine-grained temporal processes may be preserved and enable speech comprehension in the absence of environmental sound comprehension. That is, in some cases of auditory agnosia, speech may benefit from acoustic information that is less relevant for the processing of other sounds. However, the conceptual representations linked to the acoustic information of either speech and environmental sounds are likely shared. The present results corroborate this assumption.

Conclusions

For a long time, researchers considered speech to be an exceptional type of sound that activates speech- or language-specific modules (e.g., Pinker & Jackendoff, 2005). By contrasting the nature of speech with that of other sounds, researchers inferred that the brain systems and mechanisms identified for speech processing must be speech specific. Despite the apparent differences, speech and environmental sounds share various properties. Both are generated by a moving source and, thus, signify an event in the environment. Moreover, for both speech and environmental sounds, the relationship between the sound percept and the event has to be learned. Words do not mean anything unless we have learned to which concept they refer. Likewise, sounds are not associated with specific objects or events unless we have previously heard them and learned their association. For example, if we knew only electric or revolving doors and had no conceptual represen-

tation of a door handle, then we would fail to recognize the sound of a door handle being operated. Thus, like words, environmental sounds must successfully map onto a stored conceptual representation to be recognized.

Here, we outlined the processes by which these representations are mapped. We showed that initial sound processing benefits from the perceptual relationship to sonic context. If sounds activate auditory cortex, which has been only recently activated, they benefit from sensory fluidity and perceptual familiarity. Subsequently, conceptual computations emerge that are likewise facilitated by contextual overlap. Conceptual representation stored in concept-specific brain networks are activated at around 400 msec following stimulus onset and this happens more readily if they overlap with previously activated conceptual representations. In this respect, the conceptual processing of sounds is comparable to that of words and, thus, implicates a common processing system.

APPENDIX A. DESCRIPTION OF SOUNDS USED IN EXPERIMENT

<i>Prime</i>	<i>Related</i>	<i>Unrelated</i>
Entering ATM password	ATM error signal	Toy squeaking
Woman crying	Woman sniffing	Slot machine paying out
Brushing teeth	Rinsing mouth	Wrench dropping
Car stopping	Car door opening	Man shushing
Knife sharpening	Knife scrapping	Typewriter typing
Car honking	Car skidding	Metal lid dropping
Door creaking	Door closing	Siren sounding
Door knocking	Door handle pushing	Firecrackers exploding
Drum roll	Cymbals clashing	Person slurping
Faucet turning	Water gushing	Whistle blowing
Toilet flushing	Water churning	Whip cracking
Horse trotting	Horse neighing	Man gargling
Keys jingling	Door unlocking	Water bubbling
Match stick igniting	Match stick burning	Glass breaking
Motorcycle starting	Motorcycle moving off	Tennis ball bouncing
Person yawning	Person snoring	Stapler attaching
Person about to sneeze	Person sneezed	Camera photographing
Preparing to make coffee	Spoon stirring	Man choking
Sword displacing	Sword fighting	Clock ringing
Telephone dial tone	Telephone dialing	Kettle whistling
Telephone ringing	Telephone hanging up	Train horn sounding
Person gathering spit in mouth	Person clearing throat	Dog barking
Thunder roaring	Rain pouring	Sheep bleating
Truck engine starting	Truck moving	Woman yawning
Inserting coins into vending machine	Drink being dispensed	Helicopter hovering
Applying soap on hands	Washing hands	Man coughing
Woman screaming	Person running	Putting ice into glass

UNCITED REFERENCE

Özcan & van Egmond, 2009

Acknowledgments

This work was supported by the Young Investigator Award conferred to Annett Schirmer (WBS R-581-000-066-101) and an NUS AcRF grant, "Listening Strategies for New Media; Experience and Expectation," conferred to Lonce Wyse and Trevor Penney.

Reprint requests should be sent to Annett Schirmer, Department of Psychology, Faculty of Arts and Social Sciences, National University of Singapore, Block AS4, Level 2, 9 Arts Link, Singapore 117570, Singapore, or via e-mail: schirmer@nus.edu.sg.

REFERENCES

- Altmann, C. F., Doehrmann, O., & Kaiser, J. (2007). Selectivity for animal vocalizations in the human auditory cortex. *Cerebral Cortex*, *17*, 2601–2608.
- Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology*, *19*, 250–267.
- Ballas, J. A., & Mullins, T. (1991). Effects of context on identification of everyday sounds. *Human Performance*, *4*, 199–219.
- Chen, Y., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically congruent sounds modulate the identification of masked pictures. *Cognition*, *114*, 398–404.
- Chiu, C. Y. P. (2000). Specificity of auditory implicit and explicit memory: Is perceptual priming for environmental sounds exemplar specific? *Memory and Cognition*, *28*, 1126–1139.
- Chiu, C. Y. P., & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition*, *4*, 440–458.
- 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.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.
- Dick, F., Bussiere, J., & Saygin, A. P. (2002). The effects of linguistic mediation on the identification of environmental sounds. *CRL Newsletter*, *14*, 3–9.
- Dick, F., Saygin, A. P., Galati, G., Pitzalis, S., Betrovato, S., D'Amico, S., et al. (2007). What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: Evidence from functional magnetic resonance imaging and lesion data. *Journal of Cognitive Neuroscience*, *19*, 799–816.
- Farah, M. J., Hammond, K. M., Mehta, Z., & Ratcliff, G. (1989). Category-specificity and modality-specificity in semantic memory. *Neuropsychologia*, *27*, 193–200.
- Friedman, D., Cycowicz, Y. M., & Dziobek, I. (2003). Cross-form conceptual relations between sounds and words: Effects on the novelty P3. *Cognitive Brain Research*, *18*, 58–64.
- Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., et al. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage*, *40*, 1274–1286.
- Harris, J. D., Cutmore, T. R., O'Gorman, J., Finnigan, S., & Shum, D. (2009). Neurophysiological indices of perceptual object priming in the absence of explicit recognition memory. *International Journal of Psychophysiology*, *71*, 132–141.
- Holcomb, P. J., & Neville, H. (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 sentence comprehension: An fMRI study. *NeuroReport*, *12*, 1749–1752.
- Kirmse, U., Jacobsen, T., & Schröger, E. (2009). Familiarity affects environmental sound processing outside the focus of attention: An event-related potential study. *Clinical Neurophysiology*, *120*, 887–896.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*, 520–541.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience*, *7*, 302–307.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*, 203–205.
- Lenz, D., Schadow, J., Thaerig, S., Busch, N. A., & Herrmann, C. S. (2007). What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. *International Journal of Psychophysiology*, *64*, 31–38.
- Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event related potential study. *Neuroimage*, *38*, 331–345.
- Noppeney, U., Josephs, O., Hocking, J., Price, C. J., & Friston, K. J. (2008). The effect of prior visual information on recognition of speech and sounds. *Cerebral Cortex*, *18*, 598–609.
- 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.
- 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.
- Özcan, E., & van Egmond, R. (2009). The effect of visual context on the identification of ambiguous environmental sounds. *Acta Psychologica*, *131*, 110–119.
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*, 201–236.
- Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., et al. (2005). Separate neural systems for processing action or non-action-related sounds. *Neuroimage*, *24*, 852–861.
- 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.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148.
- Sable, J. J., Low, K. A., Maclin, E. L., Fabiani, M., & Gratton, G. (2004). Latent inhibition mediates N1 attenuation to repeating sounds. *Psychophysiology*, *41*, 636–642.
- Saygin, A. P., Dick, F., Wilson, S. W., Dronkers, N. F., & Bates, E. (2003). Neural resources for processing language and environmental sounds. *Brain*, *126*, 928–945.

- Saygin, A. P., Leech, R., & Dick, F. (2010). Nonverbal auditory agnosia with lesion to Wernicke's area. *Neuropsychologia*, *48*, 107–113.
- Schirmer, A. (2004). Timing speech: A review of lesion and neuroimaging findings. *Cognitive Brain Research*, *21*, 269–287.
- Schirmer, A., Escoffier, N., Li, Q. Y., Li, H., Strafford-Wilson, J., & Li, W.-I. (2008). What grabs his attention but not hers? Estrogen correlates with neurophysiological measures of vocal change detection. *Psychoneuroendocrinology*, *33*, 718–727.
- Schneider, T. R., Engel, A. K., & Debener, S. (2008). Multisensory identification of natural objects in a two-way crossmodal priming paradigm. *Experimental Psychology*, *55*, 121–132.
- Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*, 341–349.
- Schön, D., Ystad, S., Kronland-Martinet, R., & Besson, M. (2010). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience*, *22*, 1026–1035.
- Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology: Evoked Potentials*, *100*, 517–526.
- Stuart, G. P., & Jones, D. M. (1995). Priming the identification of environmental sounds. *Quarterly Journal of Experimental Psychology: A. Human Experimental Psychology*, *48*, 741–761.
- Tae, H. P. (2010). *Introduction to digital signal processing: Computer musically speaking*. Singapore: World Scientific.
- Thierry, G., Giraud, A.-L., & Price, C. (2003). Hemispheric dissociation in access to the human semantic system. *Neuron*, *38*, 499–506.
- Thompson-Schill, S. L., Aguirre, G. K., D'Esposito, M., & Farah, M. J. (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia*, *37*, 671–676.
- Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, *33*, 485–508.

Uncorrected Proof

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

During the preparation of your manuscript, the questions listed below arose. Kindly supply the necessary information.

1. Sections do not follow journal specifications. Please check.
2. Provide expanded form of ITI.
3. sd was changed to *SD*. Please verify if these changes are correct.
4. Oezcan & van Egmond, 2003; Ballas et al., 1993 were cited in the body but not in the reference list. Please check.
5. Please insert citation of Özcan & van Egmond, 2009 in the body.
6. Please insert citation of Appendix A in the body.

END OF ALL QUERIES

Uncorrected Proof