Brain and Cognition 84 (2014) 141–152

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Conceptual priming for realistic auditory scenes and for auditory words

Aline Frey [a,*], Mitsuko Aramaki [b], Mireille Besson [c,d]

[a] Laboratoire Cognitions Humaine & Artificielle, Université Paris 8, Saint-Denis, France
[b] Laboratoire de Mécanique et d'Acoustique, CNRS, Marseille, France
[c] Laboratoire de Neurosciences Cognitives, CNRS & Aix-Marseille Université, Marseille, France
[d] Cuban Neuroscience Center, Habana, Cuba

ARTICLE INFO

Article history:
Accepted 26 November 2013
Available online 30 December 2013

Keywords:
Conceptual priming
Auditory environmental scenes
Explicit and implicit processing
Event-Related Potentials (ERP)
N400

ABSTRACT

Two experiments were conducted using both behavioral and Event-Related brain Potentials methods to examine conceptual priming effects for realistic auditory scenes and for auditory words. Prime and target sounds were presented in four stimulus combinations: Sound–Sound, Word–Sound, Sound–Word and Word–Word. Within each combination, targets were conceptually related to the prime, unrelated or ambiguous. In Experiment 1, participants were asked to judge whether the primes and targets fit together (explicit task) and in Experiment 2 they had to decide whether the target was typical or ambiguous (implicit task). In both experiments and in the four stimulus combinations, reaction times and/or error rates were longer/higher and the N400 component was larger to ambiguous targets than to conceptually related targets, thereby pointing to a common conceptual system for processing auditory scenes and linguistic stimuli in both explicit and implicit tasks. However, fine-grained analyses also revealed some differences between experiments and conditions in scalp topography and duration of the priming effects possibly reflecting differences in the integration of perceptual and cognitive attributes of linguistic and nonlinguistic sounds. These results have clear implications for the building-up of virtual environments that need to convey meaning without words.

1. Introduction

Hearing someone knocking at the door typically triggers a response: "come in", hearing a klaxon in a busy street prevents crossing in potentially dangerous conditions and hearing the sound of the cork of a bottle of champagne is most often the signature of happy events. These few examples illustrate that natural auditory scenes carry some meaning that we are able to understand. The question addressed in this paper is whether the processes that allow us to understand the meaning of auditory scenes are similar or different from the processes used to attribute meaning to words. The well-known priming paradigm, in which a prime is followed by a target that is conceptually related to the prime or not (Meyer & Schvaneveldt, 1971), was used to address this issue. In most previous studies, environmental sounds have been used as primes and printed words as targets and vice versa (Orgs, Lange, Dombrowski, & Heil, 2006; Orgs, Lange, Dombrowski, & Heil, 2007; Orgs, Lange, Dombrowski, & Heil, 2008; Van Petten & Rheinfelder, 1995) or pictures were used as primes and environmental sounds as targets (Cummings et al., 2006; Plante, Van Petten, & Senkfor, 2000).

In all experiments, results showed typical priming effects with higher error rates and/or slower reaction times (RTs) for unrelated than for related targets, independently of the linguistic or nonlinguistic nature of the targets. Moreover, the N400 component of the Event-Related brain Potentials (ERPs; Kutas & Hillyard, 1980; Kutas, Lindamood, & Hillyard, 1984) that is related to the processing of meaning was larger for unrelated than for related targets whether they were words or environmental sounds (Cummings et al., 2006; Orgs et al., 2006; Orgs et al., 2007; Orgs et al., 2008; Plante et al., 2000; Van Petten & Rheinfelder, 1995). Therefore, conceptual priming effects (i.e., the differences between unrelated and related targets) were taken to be similar for words and for environmental sounds.

However, some differences were also reported in the scalp distribution of the N400 priming effect that was larger over fronto-central regions for environmental sounds and more centro-parietally distributed for words (Cummings et al., 2006; Orgs et al., 2006). Moreover, the N400 effect (i.e. the difference between unrelated and related items) shows some "paradoxical lateralization" (Van Petten & Rheinfelder, 1995), with larger amplitude over the right than the left hemisphere for words and no inter-hemispheric differences for environmental sounds (Plante et al., 2000). Similar inter-hemispheric differences in the processing of words and environmental sounds were also reported by Lebrun et al. (2001) using the Event-Related Desynchronisation

* Corresponding author.
E-mail address: aline.frey@cognition-usages.org (A. Frey).

0278-2626/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bandc.2013.11.013

(ERD) method. Moreover, in at least two studies, the N400 priming effect showed an earlier onset for environmental sounds than for words (Cummings et al., 2006; Orgs et al., 2006). Based on these similarities and differences, one may conclude that a common conceptual system is involved for processing the meaning of words and environmental sounds. This conclusion is in line with results of experiments using different brain imaging methods (MEG; fMRI) and showing that similar networks of brain regions are involved in the semantic/conceptual processing of speech and musical sounds (Koelsch, 2005; Koelsch et al., 2004; Steinbeis & Koelsch, 2008). Moreover, the pre-attentive and attentive processing of the acoustic attributes of speech and non-speech auditory items are associated with similar responses in the primary and secondary auditory cortex (Chandrasekaran, Kraus, & Wong, 2012; Chobert, Marie, François, Schön, & Besson, 2011; Marie, Kujala, & Besson, 2012; Tervaniemi et al., 2009). Similarly, Musacchia, Sams, Skoe, and Kraus (2007) showed that the brainstem evoked responses that originate from the inferior colliculus were morphologically similar for syllables and for harmonic sounds although with some differences in the amplitude and latency of the different components (see also Bidelman, Gandour, & Krishnan, 2009 and Kraus & Chandrasekaran, 2010 for an extensive review). Thus, the representations of speech and non-speech auditory items share some similarities all along the auditory pathway up to cortical brain regions. However, specific aspects of speech processing (e.g., phonological processing) may rely on brain regions, such as the middle to posterior parts of the Superior Temporal Sulcus (Hickok & Poeppel, 2007), that are less activated by the presentation of non-speech sounds. Finally, differences in the perceptual attributes of speech and non-speech sounds may also account for differences in scalp distribution and onset latency of the priming effects (Lebrun et al., 2001).

In the studies reviewed above, environmental sounds were sounds of everyday life (e.g., ringing of a telephone, dog barking) that entertained a strong relationship with the source that produced the sound. Thus, when hearing a dog barking, participants may associate the verbal label to the sound. The similarity of the N400 conceptual priming effect for words and for environmental sounds may consequently result from the use of a common linguistic encoding strategy rather than from a common conceptual system. In an attempt to minimize linguistic mediation, Schön, Ystad, Kronland-Martinet, and Besson (2010) presented sounds with not easily identifiable sources (i.e., a verbal label could not easily be associated to the sounds). Results still revealed larger N400 to targets that were unrelated to the primes. However, as words were part of the experimental design, it remained possible that participants used a linguistic strategy not only for words but also for sounds.

This problem was possibly bypassed in two recent studies that only used environmental or harmonic sounds both as primes and as targets with no words (Aramaki, Marie, Kronland-Martinet, Ystad, & Besson, 2010; Grieser-Painter & Koelsch, 2011). Aramaki et al. (2010) used a homogeneous class of impact sounds from different materials (wood, metal and glass) and an explicit category membership task. Primes were always typical impact sounds and targets were impact sounds from the same or from a different category than the primes. Moreover, to allow for a finer-grained analysis of conceptual priming, a third type of targets was presented that were ambiguous sounds. They were built by using a morphing technique that allowed interpolating the acoustic parameters characteristic of two typical sounds (e.g., wood and glass). For comparison purposes, auditory linguistic stimuli were also presented but in separate conditions (i.e., not inter-mixed with impact sounds). Words were used as primes with words, pseudo-words (i.e., pronounceable nonsense items) and nonwords (words played backward) used as targets. Results in the explicit categorization task showed highest error rates and largest N400 components for ambiguous sounds and for pseudo-words. Moreover, P300 components were largest to both typical sound targets from a different category than the primes and for nonwords. Thus, effects at the behavioral and electrophysiological levels (N400 and P300 effects) were similarly modulated by linguistic and nonlinguistic sounds. Most importantly, N400 components were elicited in a sound-sound design that minimized the possibility of linguistic mediation strategies.

Similar results were recently reported by Grieser-Painter and Koelsch (2011) using related and unrelated harmonic sounds with varying timbre and visually presented words in four stimulus combinations (Sound–Sound; Sound–Word; Word–Sound and Word–Word). When participants were asked to judge whether the prime and targets fit together (explicit task), results showed an enhanced N400 component to unrelated compared to related targets in all four stimulus combinations (no Stimulus Combination by Relatedness interaction), although the N400 effect was larger for word than for sound targets. Since words were presented visually and sounds were presented auditorilly, these differences possibly reflect a modality of presentation effect (visual *vs* auditory). By contrast, no N400 effect was found in an implicit task when participants passively listened to the same stimulus pairs in order to decide whether a probe stimulus was previously presented (recognition test). The authors concluded that, at least with short sounds (2.7 s duration on average) meaningful representations are not automatically activated in a memory task that did not require deep, elaborate sound processing. These findings thus stand in contrast with previous results showing automatic semantic priming effects for words in Lexical Decision Task (LDT; e.g., Bentin, McCarthy, & Wood, 1985; Meyer & Schvaneveldt, 1971).

In sum, both Aramaki et al. (2010) and Grieser-Painter and Koelsch (2011) have demonstrated similar conceptual priming effects for words and sounds in conditions that minimized the potential influence of linguistic mediation (sub-vocalization) strategies by using only environmental or harmonic sounds as primes and targets with no words included in the experimental conditions. Importantly, however, conceptual priming effects were only found in explicit categorization tasks and not in an implicit probe recognition task (Grieser-Painter & Koelsch, 2011). However, in a very recent study, Schirmer and collaborators found conceptual priming effects for related compared to unrelated sound pairs when participants listened to the sounds for a subsequent recognition test without explicitly focusing attention on the prime–target relationship (Schirmer, Soh, Penney, & Wyse, 2011). Thus, whether conceptual priming effects can be elicited by using implicit tasks therefore remains to be further explored.

To this aim, we conducted an experiment that included a LDT on linguistic stimuli and a modified version of the LDT on auditory environmental scenes. We also used an explicit categorization task to compare explicit and implicit conceptual priming. As in Grieser-Painter and Koelsch (2011), we used a 4×3 design to compare, within subjects and within each of the four stimulus combinations (sound–sound (SS); word–sound (WS); sound–word (SW) and word–word (WW)), the effects of three types of conceptual relationships (Related, Ambiguous, Unrelated). However, in order to minimize carry-over effects of linguistic strategies (sub-vocalization), the SS condition was always presented first and followed by the WS, SW and WW conditions. Moreover, we used not only related and unrelated words/sounds, as in Grieser-Painter and Koelsch (2011), but also ambiguous sounds and pseudo-words as targets as in Aramaki et al. (2010). Having three types of targets allowed us to compute a finer-grained analysis of conceptual priming effects. Since the N400 has been shown to be larger to pseudo-words than to unrelated auditorilly presented words (Holcomb & Neville, 1990), it was of interest to determine whether the N400