

## Familiarity of environmental sounds is used to establish auditory rules

Ursula Kirmse<sup>a,c</sup>, Erich Schröger<sup>a</sup> and Thomas Jacobsen<sup>a,b</sup>

This study addressed the extraction of long-term sound familiarity as a higher-order feature of complex, environmental sounds. Physically variable, familiar animal sounds and spectrotemporally matched, unfamiliar control sounds were presented. Participants ignored the acoustic stimuli. Infrequent deviant sounds violated the familiarity status established by the standard sounds, but no regularity on the basis of physical features. In the auditory event-related potential, deviants elicited a negative-going deflection over parietal scalp areas around 230 ms. This effect occurred for familiar deviants among unfamiliar standards and for unfamiliar deviants among familiar standards. The results indicate the establishment of an auditory regularity based on sound familiarity. This reflects the extraction of sound familiarity outside the focus of

attention.

**Keywords:** auditory event-related potential, auditory long-term memory, deviance-related processing, environmental sounds, higher-order auditory rule establishment, sound familiarity

<sup>a</sup>Institute for Psychology, Cognitive and Biological Psychology, University of Leipzig, Leipzig, <sup>b</sup>Experimental Psychology Unit, Helmut Schmidt University/University of the Federal Armed Forces Hamburg, Hamburg and <sup>c</sup>Department of Psychology, University of Konstanz, Konstanz, Germany

Correspondence to Ursula Kirmse, Department of Psychology, General Psychology, University of Konstanz, P.O. Box D36, 78457 Konstanz, Germany  
Tel: + 49 753 188 2292; fax: + 49 753 188 2971;  
e-mail: ursula.kirmse@uni-konstanz.de

### Introduction

A fast distinction of familiar and unfamiliar sounds is crucial for an appropriate evaluation of potential risks in the personal environment. Sound familiarity influences the processing of sounds even when sounds are not attended [1–11]. Familiar musical instrument sounds and trained tone patterns have been shown to elicit enhanced responses of the auditory N1 and P2 in the event-related potential (ERP) [1–3]. Recently, we have shown similar N1 and P2 enhancements for a spectrotemporally complex, familiar environmental sound [4]. These findings suggest processing differences between familiar and unfamiliar sounds already in the time range of acoustic sound analysis [12]. In the present study, we investigated whether the familiarity of complex environmental sounds is extracted and monitored outside the focus of attention.

In the auditory oddball paradigm, infrequent irregular sounds ('deviants') are presented randomly interspersed in a sequence of frequent – regular – 'standard' sounds. In the auditory ERP, deviants typically elicit deviance-related processing between 100 and 300 ms after deviance onset (usually indexed by the mismatch negativity (MMN) and the P3a; for a review, see Näätänen *et al.* [13] and Escera *et al.* [14]). This deviance-related processing indicates that the regular feature of the standard sounds has been extracted and that a deviance from this regularity has been detected [15,16]. The deviance-related processing is usually equally observed when participants ignore the acoustic stimuli. The oddball paradigm can thus be used to test which sound features are analyzed outside the focus of attention [13].

When presenting spectrotemporally complex sounds, previous oddball studies have shown enhanced deviance-related processing for auditorily familiar deviant and/or standard sounds. This has been shown for lexical words compared with pseudowords (for a review, see Pulvermüller and Shtyrov [5]), personally significant versus nonsignificant sounds (e.g. familiar voices or personal message tones [6–8]) and familiar versus unfamiliar environmental sounds [9–11,14]. Effects were observed on the MMN [5,6,9,11], the P3a [6,7,10,14], and on later processing stages related to memory recollection, semantic context integration, and the retrieval of additional stimulus information (e.g. parietal positive component/N400 [7,8,10]).

However, in these studies, deviants always violated an acoustic regularity established by the standards. The reported familiarity effects thus reflect the modulation of acoustic deviance processing by sound familiarity. In contrast, it is an open question as to whether changes in the familiarity status of complex sounds *per se* can elicit deviance-related processing when they occur without a concurrent acoustic deviancy.

Only one previous study has reported deviance-related processing that was solely related to a change in the long-term significance of auditory items. Müller-Gass *et al.* [17] presented variable sequences of a large number of acoustically variable words and linguistically matched pseudowords. Deviants only differed from standards in their lexicality (word/pseudoword). Lexical and nonlexical deviants elicited a positive ERP deflection at 165 ms and a negative deflection at 375 ms compared with the

standard ERP. Whereas the first effect possibly indicated processing related to the lexicality status *per se*, the latter effect possibly indicated the detection of the deviance in lexicality. This effect showed a parietally negative scalp distribution, which is in line with previously reported ERP effects related to sound familiarity [4,7,8]. As the participants' task was not related to the lexical status, the authors concluded that lexicality and its deviance from the context were extracted automatically.

Here, we presented long-term familiar animal sounds and spectrotemporally matched, unfamiliar control sounds. Participants ignored the sounds. Infrequent items violated the familiarity status established by the standards, but no regularity on the basis of physical features [17]. Deviance-related processing for the deviants in sound familiarity would indicate the establishment of an auditory regularity based on the familiarity of the sounds and thus the extraction of sound familiarity outside the focus of attention. We expect deviance-related processing after 200 ms from sound onset [8,17,18] and with a parietal scalp distribution [4,8,17].

## Methods

### Participants

Participants were 24 healthy adults (mean age 23.5 years, range 18–29 years, 12 women). All reported normal hearing. The experiment was performed in accordance with the Declaration of Helsinki. Participation was compensated financially or with course credits.

### Stimuli

Twelve long-term familiar sounds and 12 spectrotemporally matched, unfamiliar control sounds were presented. Familiar sounds were naturally recorded, noise-free animal vocalizations collected from various open sources on the internet [cow, cat, bird, sheep, horse, pig, frog, elephant, crow, dog, duck, lamb; duration: mean (SD) 485 (146) ms; range 201–641 ms]. Selection criteria were based on recommendations by Marcell *et al.* [19]. The frequency spectrum and the temporal envelope of each animal sound were extracted using a fast Fourier transformation/Hilbert transformation, respectively. To synthesize one control sound from each animal sound, the frequency phase information was then randomly jumbled between the frequencies present in the sound (cf. Kirmse *et al.* [4]). Frequency amplitude values were left unchanged. The resulting sound was fit to the intensity envelope of the original sound. Thus, control sounds retained the spectral and temporal structure of the animal sounds.

The representativeness of the stimuli was tested in a separate evaluation study ( $n = 8$ , cf. Kirmse *et al.* [4]). On a seven-step Likert scale (1-not familiar at all; 7-very familiar), the mean familiarity rating was 5.71 (SD 1.68) for familiar sounds and 2.51 (SD 1.49) for unfamiliar sounds [ $t(7) = 11.19$ ,  $P < 0.001$ ]. On average, participants

correctly identified 10 of the 12 familiar sounds (SD 1.07) and associated 0.38 (SD 0.52) control sounds with their original sound source [ $t(7) = 19.32$ ;  $P < 0.001$ ].

### Design and procedure

Sounds were arranged to two block types, with 12 different sounds presented per block. In deviant-familiar blocks, two familiar sounds were presented together with 10 unfamiliar control sounds. In deviant-unfamiliar blocks, two unfamiliar control sounds were presented together with 10 familiar sounds. Thus, within each block, two sounds had a deviant familiarity status and 10 sounds had a standard familiarity status (deviant probability: two of 12 = ~17%). The specific sound items presented as deviants and standards were permuted over participants. Thus, there was no fixed physical standard-deviant contrast. Deviants differed from standards regarding their familiarity status, but were no physical deviants. Each of the 12 individual sounds occurred with the same probability within the block (one of 12 = 8.3%). Familiar and unfamiliar sounds had a comparable spectral and temporal variation. For a specific deviant sound, the spectrotemporally matched control sound was equally presented within the block. This ensured that deviants did not violate the sequence by their specific length or pitch.

The present data were collected as part of a larger experiment. In the remaining part, all 12 sounds within a block were either all familiar or all unfamiliar. These data will be reported elsewhere.

Two deviant-familiar and two deviant-unfamiliar blocks were presented in a random order and intermixed with four blocks with constant familiarity. Each oddball block comprised 480 trials (40 trials per individual sound/80 deviants). The sequence of individual sounds was fully randomized within a block. By permutation across participants, all sounds contributed equally to the grand-average ERPs for deviants and standards at the group level. The difference between deviants and standards was calculated between, on average, physically identical stimuli taken from reversed oddball conditions. General differences between familiar and unfamiliar sounds thus subtract out when subtracting deviants and standards. Sounds were presented at an average level of 65 dB(A) through Sennheiser HD25.1 headphones (Sennheiser, Wedemark, Germany). Participants watched a silent, subtitled movie and ignored the sounds. A variable stimulus-onset asynchrony between 1250 and 1550 ms was used to compensate for the variable length of the sounds.

### Electroencephalogram recording and data analysis

The continuous electroencephalogram was acquired with a BioSemi ActiveTwo system (BioSemi B.V., Amsterdam, Netherlands; DC-coupled, physical low pass at 103 Hz, sampling rate 512 Hz). Data were recorded from 128 active Ag/AgCl scalp electrodes positioned radially and in

concentric cycles around Cz. Additional active electrodes recorded the signal from both mastoids, the tip of the nose, the vertical electrooculogram (superior and inferior to the right eye), and the horizontal electrooculogram (left and right outer canthus). The BioSemi-specific common mode sense electrode (CMS) was used as a reference site.

Data were offline rereferenced to the tip of the nose and were filtered with a 1–15 Hz band-pass filter (filter length 1501 data points). Trials were baseline corrected for –100 to 0 ms prestimulus electroencephalogram activity. For artifact correction, trials were excluded from averaging if voltages exceeded  $\pm 75 \mu\text{V}$  at any channel within –200 to 850 ms relative to stimulus onset. ERPs were averaged between –100 and 850 ms separately for familiar deviants, familiar standards, unfamiliar deviants, and unfamiliar standards. Distorted channels were interpolated after ERP averaging.

Deviance-related processing was analyzed between 220 and 240 ms. This time window was centered to the maximum of the global field power of the mean grand-average deviant-standard difference wave. Nine electrode positions were selected for statistical analysis that captured the scalp topography of the deviance effect: F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 (approximate correspondence in the International 10–20 system). N1 and P2 amplitudes were analyzed between 90–110 and 195–215 ms, centered around the components' peak latencies in the grand-average waveform. In addition, ERP mean values between 326 and 346 ms were subjected to statistical analysis as indicated by visual data inspection, but no relevant statistical effects were observed. Repeated-measures analysis of variance included four within-subject factors: oddball (deviant; standard), familiarity (familiar; unfamiliar), anterior electrode position (frontal – F3, Fz, F4; central – C3, Cz, C4; parietal – P3, Pz, P4), and lateral electrode position (left – F3, C3, P3; midline – Fz, Cz, Pz; right – F4, C4, P4). The Greenhouse–Geisser correction was applied where appropriate and the significance level was set to  $\alpha$  less than or equal to 0.05.

## Results

Deviants in sound familiarity, compared with standards, consistently elicited a negative deflection over the terminating slope of the P2 at parietal electrodes (Fig. 1a). Deviant-standard difference waves (Fig. 1b) showed a corresponding parietally negative peak between 220 and 250 ms. This was consistently observed for familiar deviants presented among unfamiliar standards and unfamiliar deviants presented among familiar standards (Fig. 1b).

An interaction of oddball by anterior electrode position reflected a significant deviance-related effect over parietal scalp areas [ $F(2,46) = 4.21$ ;  $\epsilon = 0.76$ ;  $P < 0.05$ ]. At parietal electrodes, significantly more negative amplitudes

were observed for deviants than standards (Fig. 2a; oddball parietal: [ $F(1,23) = 5.46$ ;  $P < 0.05$ ]). No amplitude differences were observed at frontal or central electrodes.

In line with previous studies [1–4], familiar sounds elicited an enhanced N1 compared with unfamiliar sounds (Fig. 1a; familiarity: [ $F(1,23) = 6.99$ ;  $P < 0.05$ ]). In the time window of the P2 and between 220 and 240 ms, slightly larger amplitudes for familiar sounds were observed numerically at frontolateral sites (cf. Fig. 2b). However, no effects of familiarity were significant in either time window. There were no effects of oddball on the N1 or P2. In general, no interactions of oddball  $\times$  familiarity were observed.

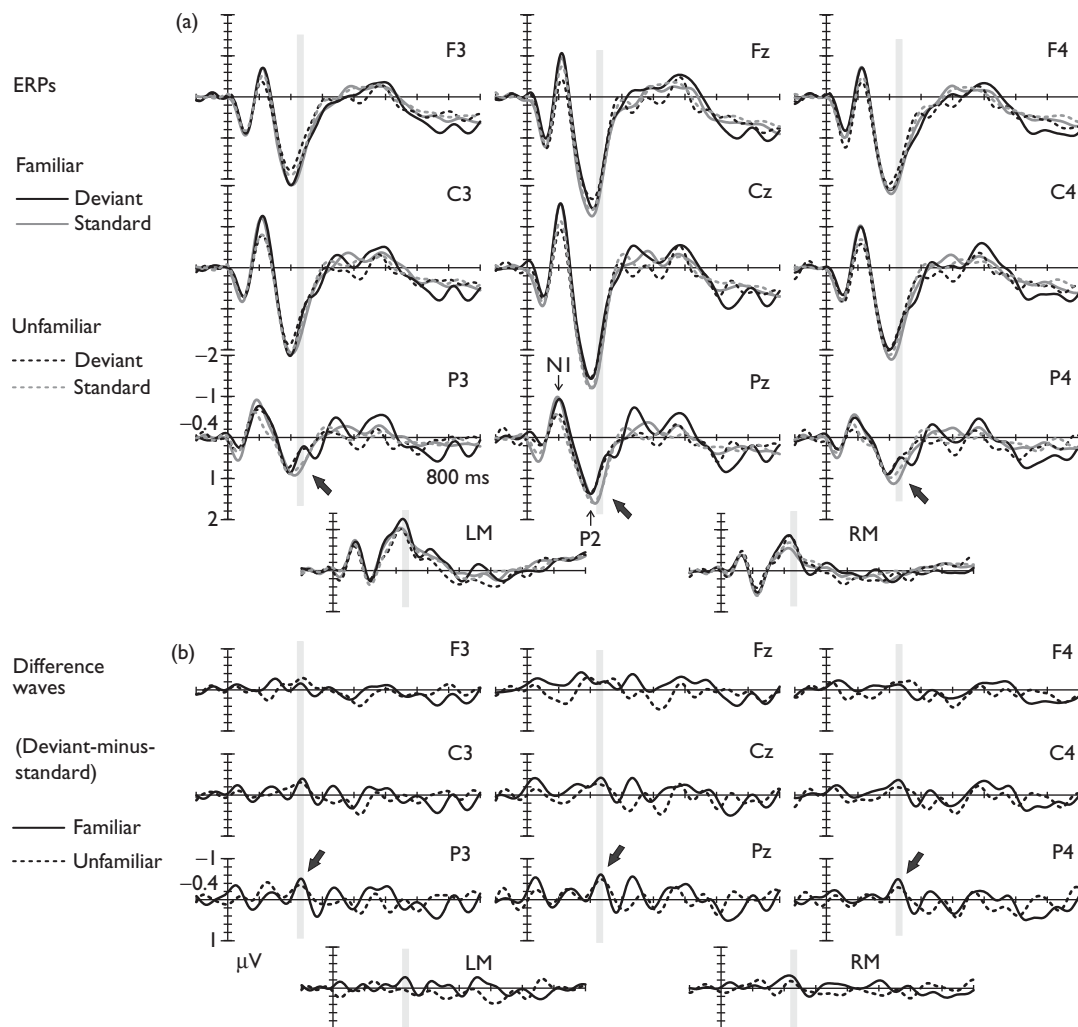
## Discussion

Environmental sound familiarity is based on stored auditory templates for the complex acoustic patterns of repeatedly encountered sounds. The present data show specific deviance-related processing for deviants in sound familiarity. This was observed even though deviants did not violate a physical rule. Deviant sounds occurred with the same probability as all other, varying sounds within the sequence. Because of the sound manipulation applied, familiar and unfamiliar sounds had a similar spectral and temporal variation. No classical, frontocentral MMN was observed. This supports the fact that familiarity deviants were indeed not processed as being acoustically deviant. Whereas previous studies [5–11] showed a modulation of acoustic deviance processing by sound familiarity, the present data indicate deviance-related processing for changes in sound familiarity *per se*.

Our results indicate that the familiarity status of the sounds, familiar versus nonfamiliar, was extracted outside the focus of attention. A rule was established for the familiarity status of the standards, against which deviances in familiarity could be detected [15,16]. The comparable deviance-related processing observed for familiar and unfamiliar deviants, among familiar/unfamiliar standards, respectively, supports the notion of a mental regularity established for sound familiarity in both block types [17]. The observed latency shortly after 200 ms is in line with Näätänen *et al.*'s [18] proposal of an integrated sound representation emerging at about 200 ms after sound onset. This is proposed to integrate the various features of the sound until this point and to crucially underlie conscious mental sound representation. Generating a 'central sound representation' may be necessary to classify spectrotemporally complex environmental sounds and to allow categorical decisions on the basis of stored auditory templates [18].

Our present findings are similar to the results of Müller-Gass *et al.* [17], who reported a parietally negative deviance effect for deviants in word lexicality. Auditory lexical words can also be considered acoustically familiar.

Fig. 1



(a) Event-related potentials (ERPs) elicited by deviants and standards for familiar and unfamiliar sounds. Gray bars mark the time window of statistical analysis. Arrows indicate the deviance-related effect. (b) Deviant-minus-standard difference waves for familiar (—) and unfamiliar (- - -) sounds (arrows indicate the deviance-related effect). A consistent deviance-related effect was observed between 220 and 250 ms at parietal electrodes. Difference waves were calculated between on group level physically identical stimuli.

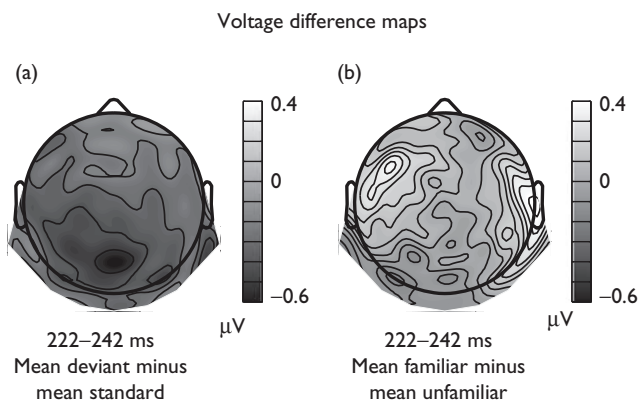
The present results suggest that the effect observed by Müller-Gass *et al.* [17] might not be language-specific. Rather, both effects may have a common basis in the existence of stored auditory patterns (cf. also Jacobsen *et al.* [9]). Similar to the present data as well, Roye *et al.* [8] observed a parietally negative deviance effect around 200 ms. This occurred for personally significant, but not personally nonsignificant, deviant mobile phone rings presented among personally nonsignificant standards.

Acoustic rule violations usually elicit frontocentral ERP effects [13]. In contrast, the observed parietal distribution of the present deviance effect is in line with previous ERP effects related to the processing of long-term sound familiarity or personal or general significance of complex

sounds [4,7,17]. The absence of a polarity inversion of this negativity at the mastoids is compatible with the hypothesis of generator locations outside the primary auditory cortex.

Consistent with previous studies [1–4], familiar sounds elicited an increased amplitude of the N1. Such initial, non-deviance-related familiarity enhancements suggest enhanced processing of sounds with stored auditory patterns directly from sound onset [1,2,4,5,12]. Such findings support modality-specific semantic memory accounts proposing a storage of modality-specific semantic representations in sensory cortical areas. These postulate possible interactions of sensory sound processing with stored auditory patterns [20,21]. General processing

Fig. 2



Voltage difference maps in the latency window of the deviance-related effect: (a) Deviant minus standard (mean for familiar and unfamiliar sounds). (b) Familiar minus unfamiliar (mean for deviants and standards).

enhancements indicate specific processing of familiar sounds. In contrast, the present data indicate that – within 200 ms after sound onset – the information about sound familiarity has been categorically extracted and has been processed up to a level that allowed the classification in rare and frequent events.

## Conclusion

The present results indicate that an auditory regularity was established based on sound familiarity. This indicates that the familiarity status of the sounds was extracted outside the focus of attention.

## Acknowledgements

U.K. was supported by the German Research Foundation (DFG) within the Research Training Group 1182, ‘Function of attention in cognition’. This research was conducted at the Unit of Cognitive and Biological Psychology, Institute of Psychology, University of Leipzig, Germany. The authors thank the reviewers and Dr Tobias Fleisch for helpful comments on the manuscript.

## Conflicts of interest

There are no conflicts of interest.

## References

- Pantev C, Roberts LE, Schulz M, Engelien A, Ross B. Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport* 2001; **12**:169–174.
- Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J Neurosci* 2003; **23**:5545–5552.
- Atienza M, Cantero JL, Dominguez-Marín E. The time course of neural changes underlying auditory perceptual learning. *Learn Mem* 2002; **9**: 138–150.
- Kirmse U, Jacobsen T, Schröger E. Familiarity affects environmental sound processing outside the focus of attention: an event-related potential study. *Clin Neurophysiol* 2009; **120**:887–896.
- Pulvermüller F, Shtyrov Y. Language outside the focus of attention: the mismatch negativity as a tool for studying higher cognitive processes. *Prog Neurobiol* 2006; **79**:49–71.
- Beauchemin M, De Beaumont L, Vannasing P, Turcotte A, Arcand C, Belin P, et al. Electrophysiological markers of voice familiarity. *Eur J Neurosci* 2006; **23**:3081–3086.
- Holečková I, Fischer C, Giard MH, Delpuech C, Morlet D. Brain responses to a subject's own name uttered by a familiar voice. *Brain Res* 2006; **1082**:142–152.
- Roye A, Jacobsen T, Schröger E. Personal significance is encoded automatically by the human brain: an event-related potential study with ringtones. *Eur J Neurosci* 2007; **26**:784–790.
- Jacobsen T, Schröger E, Winkler I, Horvath J. Familiarity affects the processing of task-irrelevant auditory deviance. *J Cogn Neurosci* 2005; **17**:1704–1713.
- Mecklinger A, Opitz B, Friederici AD. Semantic aspects of novelty detection in humans. *Neurosci Lett* 1997; **235**:65–68.
- Frangos J, Ritter W, Friedman D. Brain potentials to sexually suggestive whistles show meaning modulates the mismatch negativity. *Neuroreport* 2005; **16**:1313–1317.
- Roye A, Schröger E, Jacobsen T, Gruber T. Is my mobile ringing? Evidence for rapid processing of a personally significant sound in humans. *J Neurosci* 2010; **30**:7310–7313.
- Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. ‘Primitive intelligence’ in the auditory cortex. *Trends Neurosci* 2001; **24**:283–288.
- Escera C, Yago E, Corral MJ, Corbera S, Nunez MI. Attention capture by auditory significant stimuli: semantic analysis follows attention switching. *Eur J Neurosci* 2003; **18**:2408–2412.
- Winkler I. Interpreting the mismatch negativity. *J Psychophysiol* 2007; **21**:147–163.
- Schröger E. Mismatch negativity – a microphone into auditory memory. *J Psychophysiol* 2007; **21**:138–146.
- Müller-Gass A, Roye A, Kirmse U, Saupe K, Jacobsen T, Schröger E. Automatic detection of lexical change: an auditory event-related potential study. *Neuroreport* 2007; **18**:1747–1751.
- Näätänen R, Kujala T, Winkler I. Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology* 2011; **48**:4–22.
- Marcell MM, Borella D, Greene M, Kerr E, Rogers S. Confrontation naming of environmental sounds. *J Clin Exp Neuropsychol* 2000; **22**:830–864.
- Proverbio AM, Del Zotto M, Zani A. The emergence of semantic categorization in early visual processing: ERP indices of animal vs. artifact recognition. *BMC Neurosci* 2007; **8**:24.
- Goldberg RF, Perfetti CA, Schneider W. Perceptual knowledge retrieval activates sensory brain regions. *J Neurosci* 2006; **26**:4917–4921.