NeuroImage 188 (2019) 208–216

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Passive exposure to speech sounds modifies change detection brain responses in adults

Jari L.O. Kurkela [a,*], Jarmo A. Hämäläinen [a], Paavo H.T. Leppänen [a], Hua Shu [b], Piia Astikainen [a]

[a] Department of Psychology, University of Jyvaskyla, Jyväskylä, Finland
[b] National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

ARTICLE INFO

Keywords:
Perceptual learning
Speech sounds
Passive exposure
Event-related potentials

ABSTRACT

In early life auditory discrimination ability can be enhanced by passive sound exposure. In contrast, in adulthood passive exposure seems to be insufficient to promote discrimination ability, but this has been tested only with a single short exposure session in humans. We tested whether passive exposure to unfamiliar auditory stimuli can result in enhanced cortical discrimination ability and change detection in adult humans, and whether the possible learning effect generalizes to different stimuli. To address these issues, we exposed adult Finnish participants to Chinese lexical tones passively for 2 h per day on 4 consecutive days. Behavioral responses and the brain's event-related potentials (ERPs) were measured before and after the exposure for the same stimuli applied in the exposure phase and to sinusoidal sounds roughly mimicking the frequency contour in speech sounds. Passive exposure modulated the ERPs to speech sound changes in both ignore (mismatch negativity latency, P3a amplitude and P3a latency) and attend (P3b amplitude) test conditions, but not the behavioral responses. Furthermore, effect of passive exposure transferred to the processing of the sinusoidal sounds as indexed by the latency of the mismatch negativity. No corresponding effects in the ERPs were found in a control group that participated to the test measurements, but received no exposure to the sounds. The results show that passive exposure to foreign speech sounds in adulthood can enhance cortical discrimination ability and attention orientation toward changes in speech sounds and that the learning effect can transfer to non-speech sounds.

1. Introduction

In early infancy, cortical discrimination ability is enhanced even by passive sound exposure alone (e.g., Cheour et al., 1998; Cheour et al., 2002; Kuhl, 2004; Trainor et al., 2011). In contrast, in adulthood passive sound exposure in absence of training seems to be insufficient to affect the neural-level discrimination ability (Näätänen et al., 1993; Sheehan et al., 2005; Elmer et al., 2017) or behavioral discrimination performance (Wright et al., 2010, 2015). Instead, effects of active discrimination training have been shown in several studies by measuring the mismatch negativity (MMN) (Kraus et al., 1995; Tremblay et al., 1997, 1998; Tamminen et al., 2015), P3a (Atienza et al., 2004; Uther et al., 2006; Seppänen et al., 2012) and P2 (Atienza et al., 2002; Reinke et al., 2003; Sheehan et al., 2005) components of event-related potentials (ERPs). These components reflect pre-attentive change detection (MMN) and subsequent attention shifting (P3a) based on a memory trace formed by the learned sound feature (Näätänen et al., 2005; Polich, 2007) and sound feature encoding and stimulus classification (P2) (for a review see Crowley and Colrain, 2004).

Even though effects of passive exposure have been studied on brain responses related to pre-attentive change detection, possible effects of passive exposure on attentive change detection of sounds have not been investigated, i.e. effects on N2b and P3b components. 1-hour attentive identification training with speech sounds, however, showed learning-related changes in N2b and P3b (Alain et al., 2010). In another study, identification training resulted in only enhanced P3b responses and no changes in N2b (Ben-David et al., 2011). Similarly, attentive discrimination training with speech sounds resulted in enhanced P3b-like but not N2b-like microstates in electroencephalography (Giroud et al., 2017).

Even if previous studies have failed to demonstrate effect of passive exposure on auditory change detection in adults (Näätänen et al., 1993; Sheehan et al., 2005; Elmer et al., 2017; Wright et al., 2010, 2015), passive exposure to sounds seems not to be entirely ineffective either. Perceptual learning on an auditory discrimination task (Wright et al., 2010) or on an identification task (Wright et al., 2015) that is combined with sessions of passive exposure is more efficient than the active

* Corresponding author. Department of Psychology, Mattilanniemi 6-8, 40014 University of Jyvaskyla, Jyväskylä, Finland.
E-mail address: jari.kurkela@jyu.fi (J.L.O. Kurkela).

https://doi.org/10.1016/j.neuroimage.2018.12.010
Received 27 June 2018; Received in revised form 3 December 2018; Accepted 5 December 2018
Available online 6 December 2018

training alone as indexed by behavioral responses (Wright et al., 2010, 2015). Furthermore, passive exposure to sounds increases amplitude of the P2 component (Sheehan et al., 2005; Tremblay, 2007, 2010; Ross et al., 2013). Thus, passive exposure seems to have at least facilitating effect on auditory perceptual learning in adulthood.

One possible reason for the failure of the previous studies in demonstrating the effect of passive exposure on discrimination ability can be the short, 1–2 hour, exposure time that has been used in previous studies (Näätänen et al., 1993; Sheehan et al., 2005; Elmer et al., 2017). Active training studies have provided training over several days, and this has led to better discrimination ability as indexed by the enhancement of the MMN, P3a and P3b responses (Kraus et al., 1995; Tremblay et al., 1997; Giroud et al., 2017). Furthermore, it has been shown that sleep deprivation hinders the learning-related increase in the MMN amplitude and prevents the appearance of the P3a component (Atienza et al., 2004). Thus, the learning-related changes in cortical responses seem to be sleep-dependent, probably requiring memory consolidation during nocturnal sleep (Alain et al., 2015). Based on this assumption, it could be possible that the effects of mere passive exposure emerge if the exposure is expanded on several days, allowing memory consolidation. This has not yet been tested explicitly, however.

The evidence on generalization of the auditory learning to stimulus features not encountered during training is scarce. There are some studies showing that frequency or syllable discrimination training generalizes to closely similar untrained stimuli (for a review see Wright and Zhang, 2009). One study applied MMN to study the generalization, and showed that categorization training of labial stop consonant generalizes also to alveolar stop consonant as indicated by the shortened latency and increased amplitude of the MMN to non-trained stimuli (Tremblay et al., 1997).

In the present study, we tested two highly novel aspects of auditory perceptual learning: i) Effect of passive speech sound exposure on change detection and attention orienting in ignore and attend test conditions, and ii) if the effect of passive exposure is observed, whether it generalizes to ignored non-speech stimuli. Adult native Finnish participants were exposed to speech sounds (changes in Chinese lexical tones) for a total of 8 hours over 4 days. ERPs were recorded before and after the exposure to the same speech sounds and also to sinusoidal sounds roughly mimicking the pitch contours of the speech sounds. A control group received no exposure but participated only in the ERP recordings at the same time intervals as the experimental group.

We expect that the passive exposure would result in modulations in the ERPs, reflecting changes in both pre-attentive and attentive change detection and attention orienting toward changes (MMN, P3a, N2b, and P3b), as the exposure time is longer than in the previous studies (Näätänen et al., 1993; Sheehan et al., 2005; Elmer et al., 2017) and allows memory trace consolidation during the nights between the exposure periods (Stickgold, 2005; Alain et al., 2015). Changes in these ERP components are assumed to occur due to the formation of long-term memory representations of the sounds, making change detection and attention orienting to them more efficient (as in Näätänen et al., 1997; Winkler et al., 1999). We also hypothesized, based on the findings on sound frequency training (Wright and Zhang, 2009), that the effect of passive exposure transfers to the non-speech sounds.

2. Material and methods

2.1. Participants

A total of 39 monolingual Finnish-speaking participants (mean age = 23.0 years, standard deviation [SD] = 3.3 years; 32 females and 7 males) volunteered for the study. They were recruited with announcements in the notice boards and e-mail lists of the University of Jyväskylä. The inclusion criteria for the study were an age of 18–30 years, right-handedness, normal hearing measured using audiometry, and self-reported normal vision (or corrected to normal vision). The exclusion criteria for the study were neurological or psychiatric disorders, including sleep problems, and exposure to or training in tonal languages. However, previous exposure during trips to countries where tonal languages are spoken (maximum of 2 weeks) was accepted. Written informed consent was obtained from each participant before inclusion in the study. The experiment was undertaken in accordance with the Declaration of Helsinki, and the ethical committee of the University of Jyväskylä approved the research protocol.

The participants were divided to two groups, one group of participants were passively exposed to speech sounds (n = 18, mean age = 21.7 years, SD ± 1.7) and the other served as a control group (n = 21, mean age = 24.1 years, SD ± 3.6). Data was collected in ignore and attend test conditions (described below). From both ignore and attend conditions data of 3 participants were omitted from statistical analysis due to extensive artifacts in the EEG. After the omission, data of 18 and 21 control group participants and 18 and 15 passive exposure group participants remained for the ignore and attend test conditions, respectively. In the ignore test condition data, 66.6% of the participants in the exposure group and the same portion of the participants in the control group had some musical training or had played an instrument or sang as a hobby. In the attend test condition data, this was the case for 66.6% of the exposure group and 61.9% of the control group participants. All the participants had studied English and Swedish as a foreign language. In addition, in the ignore test condition data, 61.1% of the exposure group and 88.9% of the control group participants had studied an additional language for over 2 years. In the attend test condition data, this was the case for 73.3% of the exposure and 81.0% of the control group participants.

2.2. Stimuli

We exposed the participants to lexical tones, since Finnish belongs to a quantitative language group, and tonal changes are not part of the phonological system in this language. Therefore, we expected that training effects could be observed. Because discrimination threshold for the lexical tones applied in the study was not known for the Finnish participants, and we did not want the participants to actively listen to the sounds, two levels (large and small) of change were selected to maximize the possibility to find an exposure effect.

The sounds were prepared so that the first phoneme/a/was spoken by a female native Chinese speaker with rising (i.e., Chinese lexical tone 2) and falling (i.e., Chinese lexical tone 4) pitch contour, and they were recorded at a sampling rate of 44.1 kHz. The sounds were then digitally edited using SoundForge software (SoundForge 9, Sony Corporation, Japan) to modify them to have a duration of 200 ms. To isolate the lexical tones and keep the rest of the acoustic features identical, pitch tier transfer was performed using Praat software (Praat v5.4.06, University of Amsterdam). Pitch tier transfer generated a rising tone and a falling tone, which were identical to each other, except for a pitch contour difference in fundamental frequency (F0). These two tones were taken as the endpoint stimuli to create a continuum of lexical tones with 10 interval steps. A morphing technique was performed in MATLAB (MathWorks, Inc., MA, US), and a STRAIGHT tool (Kawahara et al., 1999) was used to create the three tones applied in the experiment. The repeatedly presented standard sound was the falling tone (Fig. 1A), and deviant sounds were a slightly falling tone (small deviant), and a rising tone (large deviant, Fig. 1A) corresponding to the tones 11, 7, and 3, respectively, on the tone continuum. All stimuli were normalized to have the same root mean square intensity. The detailed procedure concerning how the stimuli were generated was reported previously elsewhere (Xi et al., 2010).

The sinusoidal sounds were created using SoundForge software (SoundForge 9, Sony Corporation, Japan) and they had the same duration (200 ms) and start and end F0 as the corresponding speech sounds. For the standard sound, the starting frequency of F0 was 312 Hz and it gradually decreased to 180 Hz. The large deviant had the starting F0 at

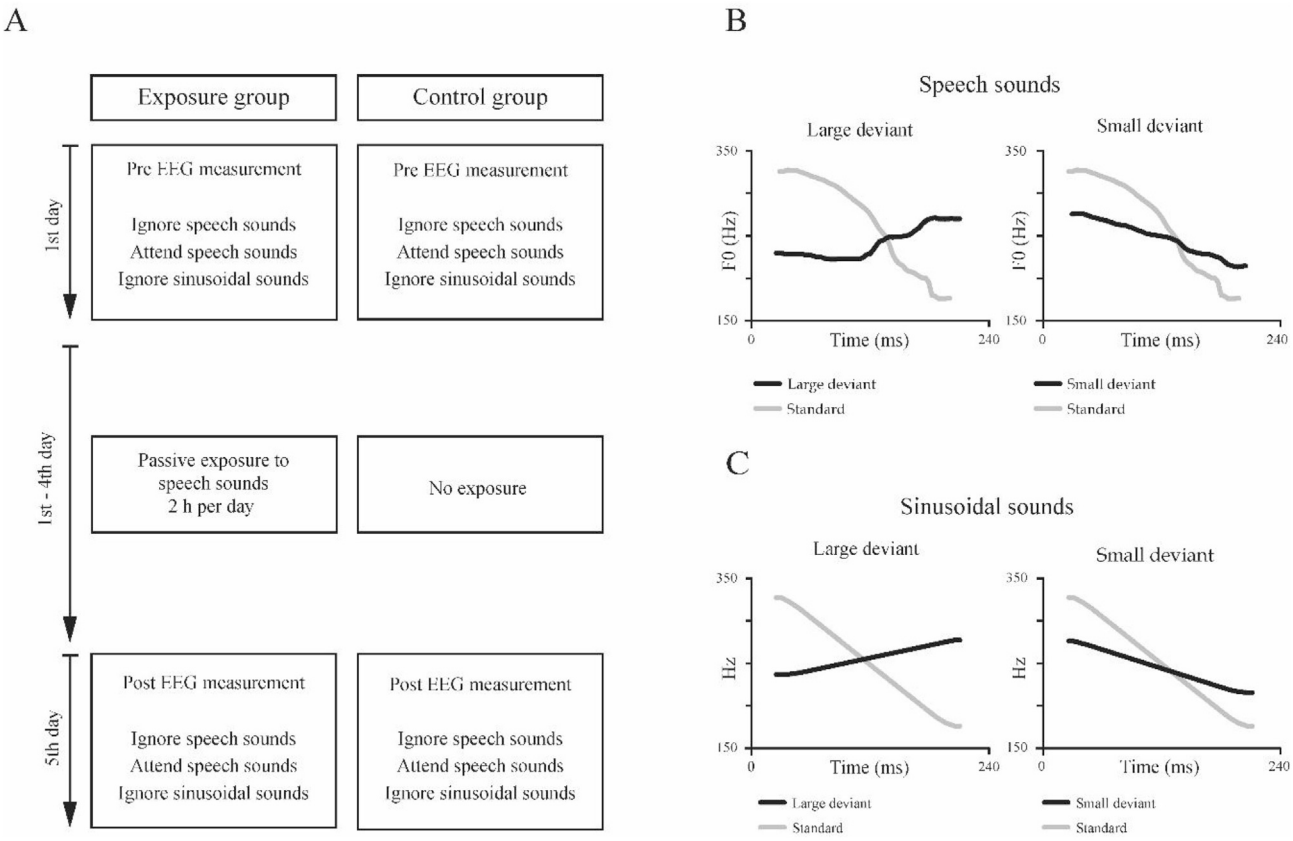

Fig. 1. Illustration of the experiment design and fundamental frequency (F0) and pitch applied in the stimuli. (A) In the pre EEG measurement, stimuli were presented in ignore and attend test conditions for both groups. After the pre EEG measurement, participants in the exposure group were exposed to speech sounds for 4 consecutive days, 2 h per day, while control group did not get any exposure. After the exposure post EEG measurement took place in the same manner as the pre EEG measurement for both groups. **(B–C)** In pre- and post-measurements, in separate stimulus blocks the stimuli were speech (phoneme/a/) or sinusoidal sounds presented in the oddball condition. The same speech sounds were also applied in the exposure phase. The black lines represent the deviant sounds (large and small deviants), and the gray lines represent the standard sounds.

233 Hz and it increased gradually to 268 Hz (Fig. 1 B). Lastly, the small deviant had a starting F0 at 268 Hz and it gradually decreased to 215 Hz (Fig. 1B).

During the pre- and post-exposure electroencephalogram (EEG) recordings and during the exposure, the sounds were presented in the oddball condition, where a frequently occurring standard stimulus (probability of 0.80) was interspersed with two deviant sounds (large or small, probability of 0.10 for each), using E-prime 1.2 (Psychology Software Tools Inc., Sharpsburg, USA) software, resulting in 1000 stimulus presentations with a sound pressure level (SPL) of 70 dB. The interstimulus interval (ISI) varied randomly between 440 and 520 ms (offset to onset). The stimuli were delivered in a pseudorandom fashion, with the restriction that consecutive deviant sounds were separated by at least two standard sounds.

2.3. Procedure

For the exposure group, experimental sessions were conducted on 5 consecutive days (Fig. 1A). On the first day, a pre-exposure EEG measurement was carried out to determine the responses to stimuli before the passive exposure. During the experiment, participants sat in a comfortable chair in a well-lit room with a video connection to the experimenter. Auditory stimuli were presented via a loudspeaker placed at approximately 50 cm above the participant's head. Two experimental conditions were applied. First, in separate stimulation blocks, we played the speech sounds (Fig. 1B) and sinusoidal sounds (Fig. 1C) to the participants and asked them to ignore the sounds and concentrate on a silent movie. Second, participants were instructed to detect changes in the speech sounds (Fig. 1B) and press a button as quickly as possible whenever they detect a deviant sound. They were not informed about the type of changes (e.g., rising or falling pitch).

From 2 to 6 hours after the pre-exposure EEG measurement, the first exposure session started. Participants watched silent movies for 2 h, while in the background the same speech sounds as used in the pre-exposure EEG measurement were presented in same oddball paradigm from the loudspeaker placed at approximately 50 cm above the participant's head with sound pressure level (SPL) of 50 dB (Fig. 1A). Participants were instructed to ignore the sounds and pay attention to the movie. Again, no information about the sounds was given to the participants. After every 30 min, a break was taken and participants were asked questions regarding the plot of the movie to keep their focus on it. The second, third, and fourth exposure sessions on the following consecutive days were the same as the first one. A total of 2 h of exposure was given on each day, summing to a total of 8 h for each participant. During the exposure, EEG time-locked to sounds was recorded, but these data are not reported here.

On day 5, a post-exposure EEG measurement was performed. The procedure was identical to that in the pre measurement (Fig. 1A).

Control participants went through the pre and post EEG measurements the same way as the exposure group on days 1 and 5, but they received no sound exposure between these measurements (Fig. 1A).

2.4. EEG measurements

Raw EEG was recorded with the Electrical Geodesics Inc. (EGI, Eugene, OR, USA) system with 128-channel sensor nets (Hydrogel GSN

128, 1.0) using Ag-AgCl electrodes. The sampling rate during the pre- and post-tests was 500 Hz, and the data were filtered online from 0.1 Hz to 200 Hz. Impedances for the electrodes were kept below 50 kΩ.

2.5. Analysis of behavioral data

Responses to the target deviant stimuli were considered hits if they occurred after the offset of the deviant sound and before the onset of the second standard stimulus onset (i.e. the hit response could occur during the post-deviant standard stimulus). The reaction times for the hits were calculated from the onset of the deviant stimulus.

2.6. EEG analysis

Offline data analysis was performed on Brain Vision Analyzer 2.1 (Brain Products GmbH). An infinite impulse response (IIR) filter for a 1-Hz low cut-off (24 dB/octave) was applied for the continuous EEG data offline. Then, independent component analysis (ICA) was performed to detect and remove eye movement artifacts from the data. Noisy channels were interpolated using a spherical spline model (Perrin et al., 1989); the average of the interpolated channels was seven. Epochs from 100 ms before to 700 ms after the stimulus onset were parsed into segments. The baseline correction was calculated based on a 100-ms prestimulus interval. Epochs including amplitude values outside the range from −200 to 200 μV, activity less than 0.5 μV, and gradients larger than 75 μV/ms were rejected within 100-ms consecutive intervals for the epoch. Epochs were averaged separately for each deviant type and the standard stimuli that preceded the deviant stimuli. The data were re-referenced offline to an average reference. A 20-Hz high cut-off filter was applied to the averaged segments. Averages that had more than 50/100 accepted epochs were included in the analysis.

2.7. Statistical analysis of behavioral data

Reaction times and discrimination accuracy were analyzed with a repeated-measures analysis of variance (ANOVA), with the deviant type (large vs. small) and session (pre vs. post) as within-subject factors and the group type as a between-subject factor (exposure vs. control).

2.8. Statistical analysis of EEG data

For the ignore condition, mean amplitude values were calculated from the time windows of 190–240 ms after stimulus onset for the MMN, of 250–300 ms for the P3a to the speech sounds, and of 300–350 ms for the P3a to the sinusoidal sounds. The time windows for the attend condition were 230–280 ms for the N2b and 360–410 ms for the P3b.

The mean values for the MMN and P3a in both speech and sinusoidal sounds in the ignore condition were extracted as a mean from three separate electrode clusters, as follows: left frontal (20, 23, 24, 28), mid-frontal (4, 11, 16, 19), and right frontal (3, 117, 118, 124), corresponding roughly to the areas of F3, Fz, and F4 in 10–20 system, respectively. For the N2b, the electrode clusters were left central (29, 30, 36, 37), mid-central (5, 6, 12, 11), and right central (87, 104, 105, 111), corresponding to C3, Fz, and C4, respectively. Finally, for the P3b, the left parietal (47, 52, 59, 60), mid-parietal (61, 62, 72, 78), and right parietal (85, 91, 92, 98) electrode clusters, corresponding to P3, Pz, and P4, respectively, were selected (Supplementary Fig. 1). The time windows and electrode clusters were selected based on the previous literature (for reviews, see Näätänen et al., 2005; Patel and Azzam, 2005; Polich, 2007) and visual inspection of the topographies and grand averaged waveforms (Supplementary Fig.s 2–5).

The latencies were analyzed from the deviant responses only since standard stimuli did not elicit clear responses for all participants. Latencies for the MMN, in both speech and sinusoidal sounds, and N2b were determined as a time point where a minimum amplitude (most negative) value for the deviant response was found from the time window of 150–260 ms and 200–310 ms, respectively. The latencies for the P3a, in both speech and sinusoidal sounds, and P3b were determined as a time point where a maximum amplitude (most positive) value for the deviant response was found from the time window of 250–350 ms and 340–460 ms, respectively. The electrode clusters for the latency analyses were the same as those applied for the amplitude analysis for each component.

Statistical analyses were performed on IBM SPSS Statistics v. 24 (IBM corporation, NY, USA). The mean amplitude values were analyzed separately for each component with a repeated-measures ANOVA with stimulus type (standard vs. deviant), deviant type (large vs. small), electrode cluster (left vs. mid vs. right) and session (pre vs. post) as within-subject factors and a between-subject factor group (exposure vs. control). The mean deviant response latency values were analyzed separately for each component with repeated-measures ANOVA, with deviant type (large vs. small), electrode cluster (left vs. mid vs. right) and session (pre vs. post) as within-subject factors and a between-subject factor group (exposure vs. control). Huynh–Feldt-corrected degrees of freedom were used whenever the sphericity assumption was violated. The corrected p-values are reported, but the degrees of freedom are reported as uncorrected. Parietal eta square (η_p^2) was used as an index of the effect size estimate. Here we give a complete report of only interaction effects that contain session x group effect since our focus is on the effect of the passive exposure. Other effects are reported in the supplementary materials.

Repeated measures of ANOVAs and paired t-tests (two-tailed with Bonferroni correction) were used to further investigate the interaction effects. For the four-way interactions, repeated measures of ANOVAs with session x group interaction effects were investigated first, and then continued with session related interactions separately for each group. For the three-way interactions session related interactions were investigated separately for each group.

P-values and confidence intervals (CIs) of 95% are reported after performing a bootstrapping with 1000 permutations. Cohen's d with pooled standard deviation was used as an index of the effect size estimate.

Whenever a statistically meaningful interaction effects of session x group was found in the ANOVA, Pearson correlation coefficients (two-tailed) were calculated between behavioral responses (reaction time, accuracy) and the corresponding ERP amplitude and latency values of the deviant responses from the post measurement. For the correlations, P-values, 99% CIs, and correlation coefficients are reported based on 1000 permutations in bootstrapping. The threshold for statistical significance was $p < 0.05$.

3. Results

3.1. Attend condition for speech sounds

3.1.1. Behavioral results

There were no interactions including session x group for the reaction times or for the accuracy of the behavioral responses. Detailed results for the behavioral responses are reported in the supplementary materials S1.1.

3.1.2. N2b component

There were no interactions including session x group for the N2b amplitude or latency (Table 1). The responses to deviant and standard sounds in N2b time window are reported in supplementary Fig. 2.

3.1.3. P3b component

For the P3b amplitude, an interaction effect of deviant type x stimulus type x session x group was found (Table 1). The following ANOVAs separately for the deviant responses (deviant type x session x group) (Supplementary Table 1) or separately for small and large deviant

Table 1
Summary of the significant effects in the repeated measures of ANOVA for the attend condition (speech sounds). * marks exposure-related effect. Degrees of freedom (df), F-values (F), P-values (P), and parietal eta squared (η_p^2) for effect sizes are reported.

Component/ variable	Effect	df	F	P	η_p^2
N2b/amplitude	Stimulus type	1,34	37.9	0.0001	0.53
	Deviant type	1,34	5.2	0.030	0.13
	Session	1,34	57.9	0.0001	0.63
	Electrode x Group	2,33	3.5	0.043	0.17
	Deviant type x Stimulus type	1,34	38.8	0.003	0.23
	Stimulus type x Electrode	2,33	6.4	0.004	0.28
	Session x Electrode	2,33	4.8	0.014	0.22
	Stimulus type x electrode x session	2,33	4.9	0.013	0.23
N2b (deviant)/ latency	Deviant type	1,34	77.36	0.0001	0.70
	Session	1,34	9.46	0.004	0.22
	Electrode	2,33	9.91	0.0001	0.38
	Deviant type x Session	1,34	4.17	0.049	0.11
P3b/amplitude	Deviant type	1,34	32.18	0.0001	0.47
	Stimulus type	1,34	97.03	0.0001	0.74
	Session	1,34	4.68	0.038	0.12
	Electrode	2,33	22.08	0.0001	0.57
	Electrode x Group	2,33	4.26	0.023	0.21
	Deviant type x Stimulus type	1,34	34.93	0.0001	0.51
	Stimulus type x Session	1,34	5.37	0.027	0.14
	Stimulus type x Electrode	2,33	33.64	0.0001	0.67
	Session x Electrode	2,33	5.05	0.012	0.23
	Stimulus type x Electrode x Group	2,33	9.36	0.001	0.36
	Deviant type x Stim type x Electrode	2,33	6.42	0.004	0.28
	Deviant type x Stim type x Session x Group*	1,34	6.50	0.015	0.16
P3b (deviant)/ latency	Deviant type	1,34	8.99	0.005	0.21
	Session	1,34	59.89	0.0001	0.64
	Electrode	1,34	4.08	0.027	0.20

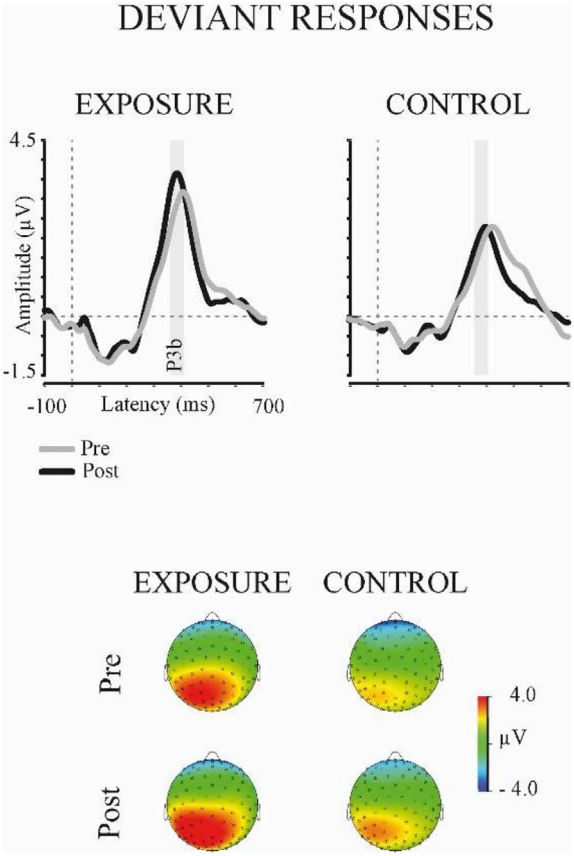

Fig. 2. Grand averaged P3b responses in the exposure group, n = 15 and the control group, n = 21. The gray lines represent responses to deviant sounds from the pre measurement, and the black lines signify the responses to deviant sounds from the post measurement. Grand averaged waveforms are presented as mean values from the collapsed electrode clusters (left, middle, and right parietal; see Supplementary Fig. 1 and 3). In lower panel, grand averaged scalp topographies of the deviant responses as a mean amplitude value of the analyzed time window at 360–410 ms from the 128 electrodes for the P3b for the exposure group and control group are shown.

responses (stimulus type x session x group) revealed no session x group interactions (Supplementary Table 2). However, the subsequent ANOVA (stimulus type x deviant type x session) performed separately for each group, revealed a significant interaction effect of stimulus type x session in the passive exposure group, $F_{1,14} = 4.97$, $p = 0.043$, $\eta_p^2 = 0.26$ (Supplementary Table 3). There was no session-related interaction effect in the control group. For the passive exposure group, subsequent t-tests were conducted where amplitude values between the pre and post measurements were compared separately for the standard and deviant sounds. These revealed that the deviant responses became significantly more positive from the pre measurement (2.98 μV ± 1.55) to post measurement (3.61 μV ± 1.20), $t(14) = 2.78$, $p = 0.032$, 95% CI [–1.09, −0.21], $d = 0.45$ (Fig. 2 and Fig. 3), but there was no change in the standard responses. There was no exposure effect for the latencies of the P3b response (Table 1).

3.1.4. Correlations between ERPs in attend condition and behavioral responses

In the exposure group, there was a marginally significant correlation between the post measurement P3b amplitude and reaction times for the small deviant, $r = -0.497$, $p = 0.059$, 99% CI [–0.80, −0.11]; the larger the response amplitude was, the faster the reaction time became. Other correlations were non-significant.

3.2. Ignore condition for speech sounds

3.2.1. MMN component

There was no exposure effect for the amplitude of the MMN response. However, for the MMM latency, an interaction effect of session x group was found (Table 2). Separate t-tests for the groups comparing the latencies between pre and post measurements showed that the deviant response latencies became shorter from the pre to post measurements in the exposure group, while no such changes in latencies were found in the control group (Figs. 4 and 5).

3.2.2. P3a component

For the amplitude of the P3a component, an interaction effect for stimulus type x session x group was found (Table 2). Subsequent ANOVA (stimulus type x session) performed separately for each group revealed that there was an interaction effect of stimulus type x session in the exposure group, $F_{1,17} = 5.66$, $p = 0.029$, $\eta_p^2 = 0.25$, while no session-related main or interaction effects were observed in the control group (Supplementary Table 4). Separate t-tests for standard and deviant responses comparing the amplitude change from pre to post measurement were conducted for the exposure group. Responses to deviant sounds increased in amplitude toward a positive polarity and the same was observed for the responses to the standard sounds (Figs. 4 and 5).

The passive exposure affected the latencies of the deviant responses in the P3a time window, as indicated by the deviant type x session x group interaction effect (Table 2). Subsequent ANOVA (deviant type x session) performed separately for each group revealed an interaction effect of

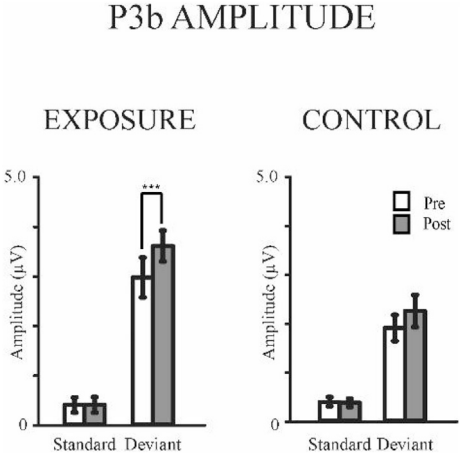

Fig. 3. Passive exposure enhanced the P3b amplitude. The mean amplitude values of P3b averaged across the three electrode clusters (left, middle, and right parietal) in 360–410 ms time window. The white bars indicate the amplitude values at the pre measurement, and the gray bars represent those at the post measurement. *** indicates a statistically significant difference ($p < 0.05$) and error bars indicate the standard error of the mean. Post exposure responses to the deviant sounds were enhanced compared to those in pre measurement. No such effect was found in the control group.

Table 2
Summary of the significant effects in the repeated measures of ANOVA for the ignore condition where the speech sounds were presented. * marks exposure-related effect. Degrees of freedom (df), F-values (F), P-values (P), and parietal eta squared (η_p^2) for effect sizes are reported.

Component/ variable	Effect	df	F	P	η_p^2
MMN/amplitude	Stimulus type	1,34	6.73	0.014	0.17
MMN (deviant)/ latency	Electrode	2,33	5.14	0.011	0.24
	Deviant type	1,34	6.92	0.013	0.17
	Deviant type x Session	1,34	4.17	0.049	0.11
	Session x Group*	1,34	6,71	0.014	0.17
P3a/amplitude	Deviant type	1,34	4,27	0.047	0.11
	Stimulus type	1,34	4.34	0.045	0.11
	Session	1,34	6.91	0.013	0.17
	Electrode	2,33	4.59	0.017	0.22
	Stimulus type x Group	1,34	46.57	<0.0001	0.58
	Session x Group*	1,34	21.12	<0.0001	0.38
	Deviant type x Group	1,34	13.76	0.001	0.29
	Dev type x Stim type x Group	1,34	12.83	0.001	0.27
	Stimulus type x Session x Group*	1,34	9.90	0.003	0.23
	Stimulus type x Electrode x Group	2,33	4.95	0.013	0.23
P3a (deviant)/ latency	Session x Group*	1,34	4.78	0.036	0.12
	Deviant type x Session x Group*	1,34	6.64	0.014	0.16

deviant type x session in the exposure group, $F_{1,17} = 9.01$, $p = 0.008$, $\eta_p^2 = 0.35$. No session-related main or interaction effects were found in the control group. Subsequent t-tests comparing latencies separately for the deviant types in the exposure group showed that after the exposure, the latency of P3a to the large change was shorter (273.6 ms ± 14.9) than it was before the exposure (294.0 ms ± 24.5), $t(17) = 2.92$, $p = 0.02$, 95% CI [8.0, 34.0], $d = 1.00$ (Fig. 5). No effect was found for the small change (Fig. 5).

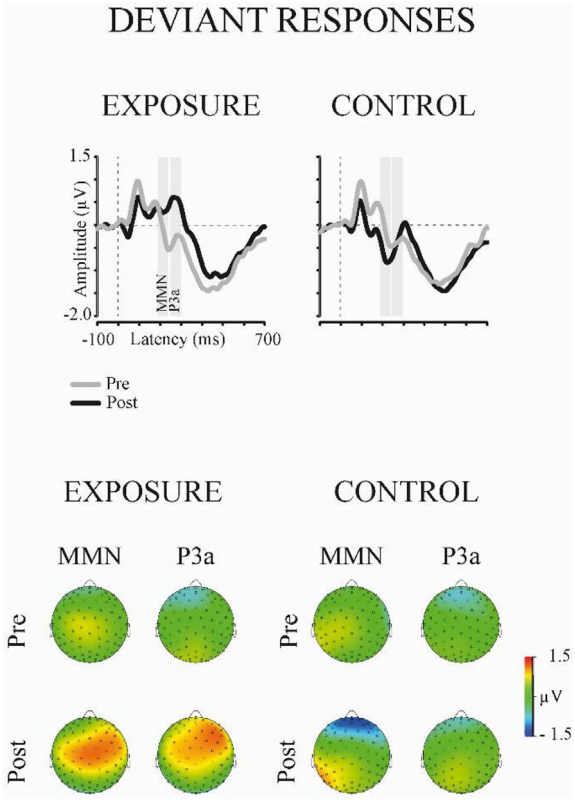

Fig. 4. Grand averaged deviant responses in the time windows of the mismatch negativity (MMN) and P3a responses in the exposure group, n = 18, and in the control group, n = 18. Grand average waveforms to deviant responses (small and large deviant averaged) are presented as mean values from the collapsed electrode clusters (left frontal, middle frontal, right frontal; see Material and methods and Supplementary Fig. 1 and 4). The gray lines represent responses in the pre measurement, while black lines represent responses in the post measurement for the exposure group and the control group. The light gray bars represent the time windows analyzed for the MMN and P3a responses (190–240 ms and 250–300 ms post stimulus onset, respectively). In lower panel, the grand averaged scalp topographies of the responses to deviant sounds as a mean amplitude value of the analyzed time window from the 128 electrodes for the MMN and P3a for the exposure group and the control group are shown.

3.3. *Ignore condition for sinusoidal sounds*

The transfer effect to non-exposed sound features was tested by presenting sinusoidal sounds roughly mimicking the pitch contours of the speech sounds in a passive oddball condition (Fig. 1C). The transfer effect was investigated for the components and variables showing group x session interaction effects in the ignore condition where speech sounds were presented, i.e. for the MMN latency, P3a amplitude and P3a latency.

3.3.1. *MMN component*

For the MMN latency, there was a significant interaction effect of electrode cluster x session x group (Table 3). Subsequent ANOVA (electrode cluster x session) performed separately for each group revealed an significant main effect session for both groups $F_{1,17} = 15.03$, $p = 0.001$, $\eta_p^2 = 0.47$; $F_{1,17} = 5.90$, $p = 0.027$, $\eta_p^2 = 0.26$ respectively (Supplementary table 5). Post hoc paired samples t-tests comparing latencies between pre and post measurements separately for groups showed that in the exposure group the latencies got significantly shorter from pre measurement (234.9 ms ± 10.58) to post measurement (215.27 ms ± 21.44), $t(17) = 3.88$, $p = 0.014$, 95% CI [10.43, 29.74], $d = 1.16$) (see Fig. 6). There were no changes in the latencies in the control group.

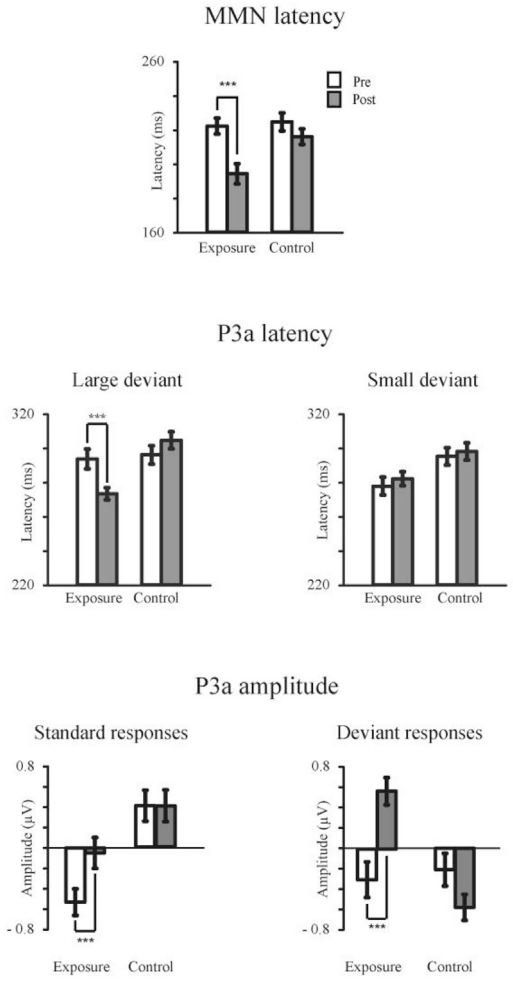

Fig. 5. Mismatch negativity latency and P3a latency and amplitude are enhanced after the passive exposure. The white and gray bars represent the mean values from the pre and post measurements, respectively. Error bars indicate the standard error of the mean. *** indicates a statistically significant ($p < 0.05$) difference. The latency of the response to the deviant sounds in the MMN time window shortened from the pre measurement (222.45 ms ± 19.61) to the post measurement (194.57 ms ± 24.81), $t(17) = 4.79$, $p = 0.00$, 95% CI [17.44, 39.78], $d = 1.25$), in the exposure group, while there were no changes in the control group. In addition, in the P3a time window, deviant response latencies became significantly shorter for the large deviant after the exposure, but not for the small deviant, and there were no changes in the control group. Standard response amplitudes were more positive after the exposure (−0.05 μV ± 0.64) compared to the pre measurement (−0.53 μV ± 0.55), $t(17) = 3.18$, $p = 0.026$, 95% CI [–0.79, −0.19], $d = 0.80$. Also deviant response amplitudes became more positive after the exposure (0.57 μV ± 0.57) compared to the pre measurement (−0.30 μV ± 0.74), $t(17) = 4.70$, $p = 0.002$, 95% CI [–1.23, −0.53], $d = 1.32$.

3.3.2. P3a component

For the P3a amplitude, there was an interaction effect of stimulus type x electrode cluster x session x group (Table 3). The following ANOVAs (electrode cluster x session x group) for the deviant responses (Supplementary table 6) or (electrode cluster x session x group) performed separately for each electrode cluster revealed no session x group interaction effects (Supplementary table 7) Subsequent ANOVA (stimulus type x electrode cluster x session) performed separately for each group revealed significant interaction effect of stimulus type x electrode cluster x session in the control group, while there was no session-related effects in the exposure group (Supplementary Table 8). T-tests comparing deviant and standard responses separately from pre measurement to post measurement in each electrode cluster was performed in the control group. They did not reveal any significant results.

Table 3
Summary of the significant effects in the repeated measures of ANOVA for the sinusoidal sounds in the ignore condition. * marks exposure-related effect. Degrees of freedom (df), F-values (F), *P*-values (*P*), and Partial eta squared (η_p^2) for effect size are reported.

Component/ variable	Effect	df	F	*P*	η_p^2
MMN (deviant)/ Latency	Electrode	2,33	6.89	0.003	0.29
	Session x Group*	1,34	19.57	<0.001	0.37
	Session x Electrode x Group*	2,33	4.52	0.018	0.22
P3a/Amplitude	Stimulus	1,34	37.9	<0.001	0.53
	Electrode	2,33	8.4	0.001	0.34
	Stimulus x Group	1,34	12.4	0.001	0.27
	Stimulus x Session	1,34	4.2	0.047	0.11
	Stimulus x Session x Electrode x Group*	2.33	4.0	0.027	0.20
P3a (deviant)/ latency	Deviant type	1,34	7.36	0.01	0.18
	Deviant type x Group	1,34	5.65	0.023	0.14

4. Discussion

Here we show in adult humans that passive exposure to foreign speech sounds for 4 consecutive days, 2 h per day, enhanced the neural discrimination ability and attention orientation toward changes in the speech sounds as indexed by ERPs recorded in ignore and attend test conditions. The effect of passive exposure to auditory change detection mechanism has earlier been found only in infants (Cheour et al., 1998, 2002; Kuhl, 2004; Trainor et al., 2011). In the attend test condition, effect of passive exposure was demonstrated as enhanced P3b amplitude. In the ignore test condition, effects of passive exposure were demonstrated as shortened latency of the MMN and enhanced amplitude and shortened latency of the P3a.

The learning effect generalized to some extend to novel sounds: the latency of the MMN shortened to the sinusoidal sounds not encountered during the exposure phase. This effect was demonstrated only in the exposure group, not in the control group.

Effects of auditory perceptual learning have rarely been tested for attentive change detection. Here we showed that passive exposure enhanced the amplitude of the P3b and there was trend towards significant correlation between the enhanced P3b amplitude and shortened behavioral reaction times to small deviant. Previously, it has been shown that when the perceptual task becomes easier, the P3b amplitude increases (for review see Polich, 2007). Our results are also in line with one previous study which showed that active training to discriminate speech sounds enhances the microstates related to the P3b component accompanied by improvements in behavioral reaction times (Giroud et al., 2017). In the light of the context-updating model (Polich, 2007), passive exposure seems to ease comparison process between the representation of the standard sound in memory and the deviant sound input, which is also reflected as shortened reaction times.

N2b was the other component that was investigated in the attend test condition. Here, the amplitude of the N2b was not enhanced, nor was its latency shortened due to passive exposure. In prior studies applying attentive training, in line with our findings, N2b was not enhanced during identification task (Ben-David et al., 2011) or during discrimination task (Giroud et al., 2017). However, a study that had longer practice period than in study by Ben-David et al. (2011) reported that enhancement in ability to identify speech sounds were followed by increased N2b (Alain et al., 2010). It remains thus unclear whether perceptual learning requires attentive training to modulate the N2b, and if so, whether the training should be identification training, instead of discrimination training.

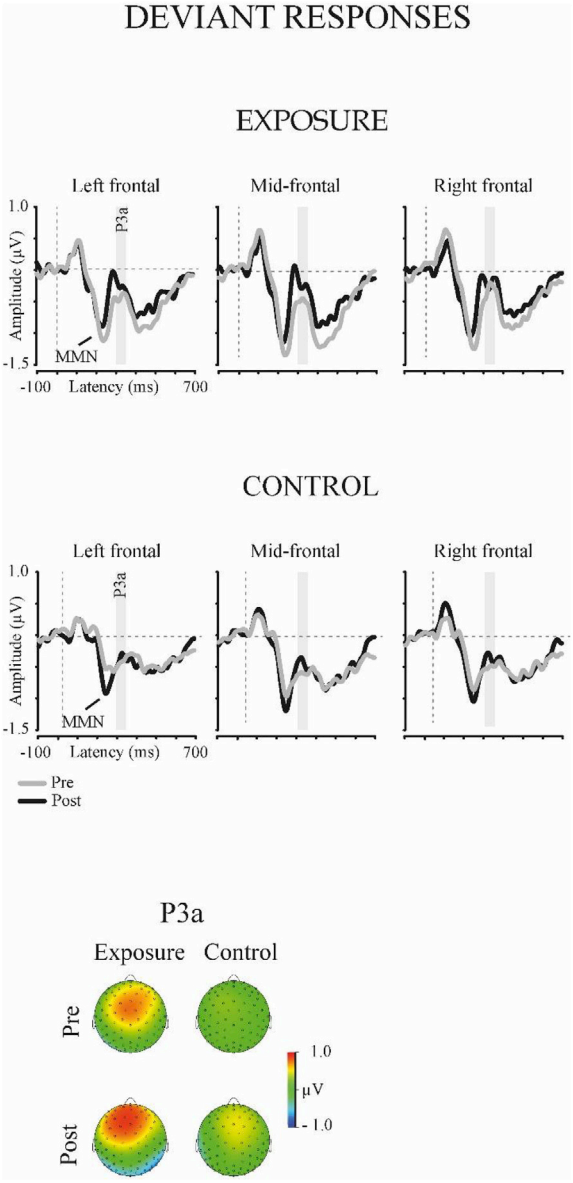

Fig. 6. Grand averaged MMN and P3a responses to sinusoidal sounds. Grand averaged waveforms of responses reflecting MMN and P3a in the exposure group (n = 18) and the control group (n = 18). The gray lines represent responses to deviant sounds (large and small deviant types averaged) from the pre measurement, while black lines represent responses to deviant sounds from the post measurement for the exposure group and the control group. The light gray bars represent the time windows applied in the analysis of for the P3a amplitude (mean amplitude value between 300 and 350 ms). The mean scalp topographies of the differential response (standard subtracted from the deviant, deviant types averaged) from the 128 electrodes for the P3a for the exposure group and the control group from the analyzed time window. Although there was a significant stimulus type x electrode cluster x session x group effect for the P3a amplitude, post-hoc tests did not reveal any generalization of the exposure effect. Please note that the MMN amplitude was not investigated since there was no exposure effect for it.

No exposure-related effects were found for behavioral responses measured in the attentive test condition. Reaction times decreased from pre to post measurement in both groups, and the decrease was larger for the small deviant than for the large deviant. The detection of the large deviant was more accurate (97.0%) than the detection of small deviant (88.4%) but there was no change from pre to post test in accuracy. Sometimes neural changes related to auditory learning precede its behavioral indices (Tremblay et al., 1998). It is possible that the passive exposure applied here should have been more extended in order to induce changes at behavioral level.

In the ignore test condition, the peak latency of the MMN response to the speech sounds was significantly shorter after the exposure compared with that before the exposure, while no latency changes were observed in the control group. Changes in MMN latency have been interpreted as enhanced discrimination ability due to enhanced memory traces of deviant sounds in studies where attentive training has been found to modulate the latency of MMN (Kraus et al., 1995; Tremblay et al., 1997, 1998). Our results show that similar modulation of the MMN can be induced by mere passive exposure. However, the MMN amplitude did not show any changes due to the passive exposure, although it typically increases in studies utilizing active training (Näätänen et al., 1993; Kraus et al., 1995; Tremblay et al., 1997). Also in our previous study in rats, 36-h passive exposure to spectro-temporal changes in speech sound /a/ enhanced the mismatch response amplitude (Kurkela et al., 2016).

Even though the amplitude of the MMN response was not changed due to passive exposure, the following P3a component's amplitude was enlarged. Also the latency of P3a was shortened, for the large deviant. Previously, active training of discriminating pitch (Seppänen et al., 2012), tone sequences (Atienza et al., 2004) or to learning to use Morse code (Uther et al., 2006) have led to increased P3a amplitude. Our results demonstrate that mere passive exposure suffices changes in the involuntary attention-shifting mechanism which the P3a is typically linked to.

Here passive exposure expanded over four consecutive days, totaling in eight hours of exposure. Studies applying attentive training have usually also spread the training sessions over several days (6–9 days) (for example, Kraus et al., 1995; Tremblay et al., 1997). Previous studies of passive exposure have applied only short exposure session during a single day, and in these works, no effects of passive exposure in ERPs reflecting sound discrimination have been demonstrated (Näätänen et al., 1993; Sheehan et al., 2005; Elmer et al., 2017). It can be assumed that we were able to demonstrate the effect of passive exposure in the ERPs related to change detection and attention shifting because the exposure was relatively long-lasting. In addition, the exposure that was extended for four consecutive days could have facilitated memory consolidation for the exposed sounds during the nocturnal sleep, allowing the emergence of the exposure effect. Previous studies applying attentive training have shown that sleep deprivation prevents the emergence or enhancement of the MMN and the P3a (Alain et al., 2015; Atienza et al., 2004). It thus seems that nocturnal sleep is a crucial factor for the emergence of learning-related enhancement in the change detection, probably due to memory consolidation for the learned sounds.

We also investigated the generalization of the perceptual learning from speech sounds to sinusoidal sounds. This was studied for the MMN latency and P3a latency and amplitude since these showed the effect of passive exposure. For the sounds that were mimicking the pitch contours of the speech sounds, the peak latency of the deviant responses in the MMN time window was significantly shorter in the exposure group after the exposure than before the exposure. No such effect was found in the control group. This pattern of results can be interpreted to reflect transfer of learning at a neural level due to passive exposure to speech sounds. Our result related to generalization is in line with the results of previous behavioral studies showing that learning to discriminate sound frequencies or syllables generalizes to closely similar sounds (for a review, see Wright and Zhang, 2009). Furthermore, our findings are in line with a study, where attentive training induced changes in the MMN (Tremblay et al., 1997). The learning effect also transferred to novel speech stimuli, i.e. from one place of articulation (labial) to another (alveolar). Our results extend this finding by showing that similar transfer effect can be induced by mere passive exposure and from speech to non-speech sounds.

The generalization effect was not observed for the P3a amplitude or latency. It thus seems that in the case of passive exposure latency changes are more sensitive than amplitude changes to reflect generalization

effect.

In summary, passive exposure to foreign speech sounds for 2 h for 4 consecutive days induced plastic cortical changes related to change detection and attention shift mechanisms. As indexed by ERPs, this was demonstrated in the attend test condition by increased P3b amplitude and in the ignore test condition by the shortened latency of the MMN and P3a as well as increased amplitude of the P3a component. In addition, the latency of the MMN shortened to the sinusoidal sounds not encountered during the exposure, reflecting generalization of the learning effect. For the first time, these results demonstrated that mere passive exposure to sounds can induce plastic changes related to change detection in the adult human brain, which was previously thought to happen only in infancy during the sensitive period. Changes in brain responses occurred from 8 h of exposure. This encourages testing the effectiveness of passive exposure in real-life language learning situations.

Disclosure statement

The authors report no conflicts of interest.

Acknowledgements

The study was supported by the Academy of Finland (project number 273134 for Piia Astikainen) and by the Natural Science Foundation of China (grant number 81461130018 for Hua Shu).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2018.12.010.

References

Ben-David, B.M., Campeanu, S., Tremblay, K.L., Alain, C., 2011. Auditory evoked potentials dissociate rapid perceptual learning from task repetition without learning. Psychophysiology 48, 797–807. https://doi.org/10.1111/j.1469-8986.2010.01139.x.

Alain, C., Campeanu, S., Tremblay, K., 2010. Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. J. Cognit. Neurosci. 22, 392–403. https://doi.org/10.1162/jocn.2009.21279.

Alain, C., Da Zhu, K., He, Y., Ross, B., 2015. Sleep-dependent neuroplastic changes during auditory perceptual learning. Neurobiol. Learn. Mem. 118, 133–142. https://doi.org/10.1016/j.nlm.2014.12.001.

Atienza, M., Cantero, J.L., Dominguez-Marin, E., 2002. The time course of neural changes underlying auditory perceptual learning. Learn. Mem. 9, 138–150. https://doi.org/10.1101/lm.46502.

Atienza, M., Cantero, J.L., Stickgold, R., 2004. Posttraining sleep enhances automaticity in perceptual discrimination. J. Cogn. Neurosci. 16, 53–64. https://doi.org/10.1162/089892904322755557.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., Näätänen, R., 1998. Development of language-specific phoneme representations in the infant brain. Nat. Neurosci. 1, 351–353. https://doi.org/10.1038/1561.

Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., Raz, A., Kaipio, M.L., Hiltunen, J., Aaltonen, O., Savela, J., Hämäläinen, H., 2002. Speech sounds learned by sleeping newborns. Nature 415, 599–600. https://doi.org/10.1038/415599b.

Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115, 732–744. https://doi.org/10.1016/j.clinph.2003.11.021.

Elmer, S., Hausheer, M., Albrecht, J., Kühnis, J., 2017. Human brainstem exhibits higher sensitivity and specificity than auditory-related cortex to short-term phonetic discrimination learning. Sci. Rep. 7, 7455. https://doi.org/10.1038/s41598-017-07426-y.

Giroud, N., Lemke, U., Reich, P., Matthes, K.L., Meyer, M., 2017. Longitudinal auditory learning facilitates auditory cognition as revealed by microstate analysis. Biol. Psychol. 123, 25–36. https://doi.org/10.1016/j.biopsycho.2016.11.007.

Kawahara, H., Masuda-Katsuse, I., de Cheveigne, A., 1999. Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds. Speech Commun. 27, 187–207. https://doi.org/10.1016/S0167-6393(98)00085-5.

Kraus, N., McGee, T., Carrell, T.D., King, C., Tremblay, K., Nicol, T., 1995. Central auditory system plasticity associated with speech discrimination training. Cognit. Neurosci. 7, 25–32. https://doi.org/10.1162/jocn.1995.7.1.25.

Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. Nat. Rev. Neurosci. 5, 831–843. https://doi.org/10.1038/nrn1533.

Kurkela, J.L.O., Lipponen, A., Hämäläinen, J.A., Näätänen, R., Astikainen, P., 2016. Passive exposure to speech sounds induces long-term memory representations in the auditory cortex of adult rats. Sci. Rep. 6 https://doi.org/10.1038/srep38904, 38904.

Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., Paavilainen, P., 1993. Development of a memory trace for a complex sound in the human brain. Neuroreport 4 (5), 503–506.

Näätänen, R., et al., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 385. https://doi.org/10.1038/385432a0, 432.

Näätänen, R., Jacobsen, T., Winkler, I., 2005. Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. Psychophysiology 42, 25–32. https://doi.org/10.1111/j.1469-8986.2005.00256.x.

Patel, S.H., Azzam, P.N., 2005. Characterization of N200 and P300: selected studies of the event-related potential. Int. J. Med. Sci. 2, 147–154. https://doi.org/10.7150/ijms.2.147.

Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. Electroencephalogr. Clin. Neurophysiol. 72, 184–187. https://doi.org/10.1016/0013-4694(89)90180-6.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019.

Reinke, K.S., He, Y., Wang, C., Alain, C., 2003. Perceptual learning modulates sensory evoked response during vowel segregation. Cognit. Brain Res. 17, 781–791. https://doi.org/10.1016/S0926-6410(03)00202-7.

Ross, B., Jamali, S., Tremblay, K.L., 2013. Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. BMC Neurosci. 14, 151. https://doi.org/10.1186/1471-2202-14-151.

Seppänen, M., Pesonen, A.K., Tervaniemi, M., 2012. Music training enhances the rapid plasticity of P3a/P3b event-related brain potentials for unattended and attended target sounds. Atten. Percept. Psychophys. 74, 600–612. https://doi.org/10.3758/s13414-011-0257-9.

Sheehan, K.A., McArthur, G.M., Bishop, D.V.M., 2005. Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? Cognit. Brain Res. 25, 547–553. https://doi.org/10.1016/j.cogbrainres.2005.08.007.

Stickgold, R., 2005. Sleep-dependent memory consolidation. Nature 437, 1272–1278. https://doi.org/10.1038/nature04286.

Tamminen, H., Peltola, M.S., Kujala, T., Näätänen, R., 2015. Phonetic training and non-native speech perception—new memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. Int. J. Psychophysiol. 97, 23–29. https://doi.org/10.1016/j.ijpsycho.2015.04.020.

Trainor, L.J., Lee, K., Bosnyak, D.J., 2011. Cortical plasticity in 4-month-old infants: specific effects of experience with musical timbres. Brain Topogr. 24, 192. https://doi.org/10.1007/s10548-011-0177-y.

Tremblay, K., 2007. Training-related changes in the brain: evidence from human auditory-evoked potentials. Semin. Hear. 28 (2), 120–132. https://doi.org/10.1055/s-2007-973438.

Tremblay, K., Kraus, N., Carrell, T.D., McGee, T., 1997. Central auditory system plasticity: generalization to novel stimuli following listening training. J. Acoust. Soc. Am. 102 (6), 3762–3773. https://doi.org/10.1121/1.420139.

Tremblay, K., Kraus, N., McGee, T., 1998. The time course of auditory perceptual learning: neurophysiological changes during speech-sound training. Neuroreport 9, 3557–3560. https://doi.org/10.1097/00001756-199811160-00003.

Tremblay, K.L., Inoue, K., McClannahan, K., Ross, B., 2010. Repeated stimulus exposure alters the way sound is encoded in the human brain. PLoS One 5. https://doi.org/10.1371/journal.pone.0010283 e10283.

Uther, M., Kujala, A., Huotilainen, M., Shtyrov, Y., Näätänen, R., 2006. Training in Morse code enhances involuntary attentional switching to acoustic frequency: evidence from ERPs. Brain Res. 1073, 417–424. https://doi.org/10.1016/j.brainres.2005.12.047.

Winkler, et al., 1999. Brain responses reveal the learning of foreign language phonemes. Psychophysiology 36, 638–642. https://doi.org/10.1111/1469-8986.3650638.

Wright, B.A., Zhang, Y., 2009. A review of the generalization of auditory learning. Philos. Trans. R. Soc. Lond. B Biol. Sci. 364, 301–311. https://doi.org/10.1098/rstb.2008.0262.

Wright, B.A., Sabin, A.T., Zhang, Y., Marrone, N., Fitzgerald, M.B., 2010. Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. J. Neurosci. 30, 12868–12877. https://doi.org/10.1523/JNEUROSCI.0487-10.2010.

Wright, B.A., Baese-Berk, M.M., Marrone, N., Bradlow, A.R., 2015. Enhancing speech learning by combining task practice with periods of stimulus exposure without practice. J. Acoust. Soc. Am. 138, 928–937. https://doi.org/10.1121/1.4927411.

Xi, J., Zhang, L., Shu, H., Zhang, Y., Li, P., 2010. Categorical perception of lexical tones in Chinese revealed by mismatch negativity. Neuroscience 170, 223–231. https://doi.org/10.1016/j.neuroscience.2010.06.077.