



HAL
open science

Effects of Psychoacoustic Training on the Pre-Attentive Processing of Harmonic Sounds and Syllables

Aline Frey, Mylène Barbaroux, Eva Dittinger, Mireille Besson

► **To cite this version:**

Aline Frey, Mylène Barbaroux, Eva Dittinger, Mireille Besson. Effects of Psychoacoustic Training on the Pre-Attentive Processing of Harmonic Sounds and Syllables. *Journal of Speech, Language, and Hearing Research*, 2022, 65, pp.2003 - 2015. 10.1044/2022_jslhr-21-00441 . hal-03774126

HAL Id: hal-03774126

<https://hal.science/hal-03774126v1>

Submitted on 9 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Research Article

Effects of Psychoacoustic Training on the Pre-Attentive Processing of Harmonic Sounds and Syllables

Aline Frey,^{a,b} Mylène Barbaroux,^a Eva Dittinger,^{a,c} and Mireille Besson^{a,b} 

^aLaboratoire de Neurosciences Cognitives (LNC), CNRS, and Aix-Marseille University, Université Publique de France ^bInstitute for Language and Communication in the Brain (ILCB), Aix-Marseille University, France ^cLaboratoire Parole et Langage (LPL), CNRS, and Aix-Marseille University, Université Publique de France

ARTICLE INFO

Article History:

Received August 14, 2021

Revision received December 30, 2021

Accepted February 11, 2022

Editor-in-Chief: Peggy B. Nelson

Editor: Christian E. Stipf

https://doi.org/10.1044/2022_JSLHR-21-00441

ABSTRACT

Purpose: This article aimed at investigating the neural underpinnings of music-to-language transfer effects at the pre-attentive level of processing.

Method: We conducted a longitudinal experiment with a test–training–retest procedure. Nonmusician adults were trained either on frequency (experimental group) or on intensity (control group) of harmonic tones using methods from psychophysics. Pre- and posttraining, we recorded brain electrical activity and we analyzed the mismatch negativity (MMN) and the P3a component both to harmonic complex sounds and to syllables varying in frequency.

Results: Frequency training influenced the pre-attentive perception of pitch for large harmonic deviant sounds but not for syllables.

Conclusion: Results are discussed in terms of near and far transfer effects from psychoacoustic training to pre-attentive pitch processing and as possibly showing some limits to transfer effects.

Musicians typically show improved auditory acuity compared with nonmusicians regarding the perception of sounds varying in frequency (Micheyl et al., 2006; Spiegel & Watson, 1984; Tervaniemi et al., 2005) and in duration (Jeon & Fricke, 1997). For instance, in seminal studies, Micheyl et al. (2006) showed that pitch discrimination thresholds were 6 times lower for musicians than for nonmusicians, and Tervaniemi et al. (2005) demonstrated that musicians detected rare deviant sounds that were 0.8% or 2% higher in frequency than standard sounds better than nonmusicians. There is also evidence, based on behavioral and/or neural measures, that improved auditory processing in musicians facilitates the processing of speech sounds (see Besson et al., 2017, for a review Anvari et al., 2002; Bidelman & Alain, 2015; Delogu et al., 2010; Elmer et al., 2012; Lima & Castro, 2011; Marie et al., 2011; Wong et al., 2007). Two different and not mutually exclusive interpretations, the cascade and

multidimensional interpretations (Besson et al., 2011, 2017) have been proposed to explain transfer effects from musical expertise to different levels of speech processing. The first interpretation is based on enhanced bottom-up processing in musicians compared with nonmusicians so that enhanced auditory perception in musicians drives higher level of performance in speech perception and comprehension. This is the interpretation that is tested in the present experiment. The second interpretation is based on enhanced top-down processes in musicians compared with nonmusicians, so that general cognitive abilities necessary for speech processing, such as attention (Strait et al., 2010), short- and long-term memory (Dittinger et al., 2021; George & Coch, 2011), and executive functions (Moradzadeh et al., 2015), are enhanced in musicians. At the neural level, transfer effects possibly reflect shared neural resources for language and music (Patel, 2011).

However, most studies conducted so far are cross-sectional studies testing for correlations between musical expertise and speech processing, and as pointed out by Schellenberg (2004), correlation is not causality. Cross-sectional studies do not establish that music training is the

Correspondence to Mireille Besson: mireille.besson@univ-amu.fr. **Disclosure:** The authors have declared that no competing financial or non-financial interests existed at the time of publication.

cause of improved speech processing because between-groups differences other than music training can explain the results (e.g., pre-existing genetic differences, motivation, and sociocultural background). Demonstrating causality requires longitudinal studies, in which music-to-language transfer effects are evidenced for nonmusicians trained with music during the course of the experiment. Overall, results showed that relatively short-term music training, ranging from a few weeks to 2 or 3 years, has a positive influence on several aspects of speech processing at both the behavioral and neural levels (see Besson et al., 2017, for review; Degé & Schwarzer, 2011; Flaugnacco et al., 2015; Frey et al., 2019; Gordon et al., 2015; Habib et al., 2016; Jentschke & Koelsch, 2009; Linnavalli et al., 2018; Moreno et al., 2015; Nan et al., 2018). Therefore, these longitudinal studies, together with experiments in which mediation analyses, are used to test for causality (Barbaroux et al., 2021; Baron & Kenny, 1986; Foncubiarta et al., 2020; Gomez-Dominguez et al., 2019), provided strong evidence that music training is the cause of the facilitatory effects observed at various levels of speech processing.

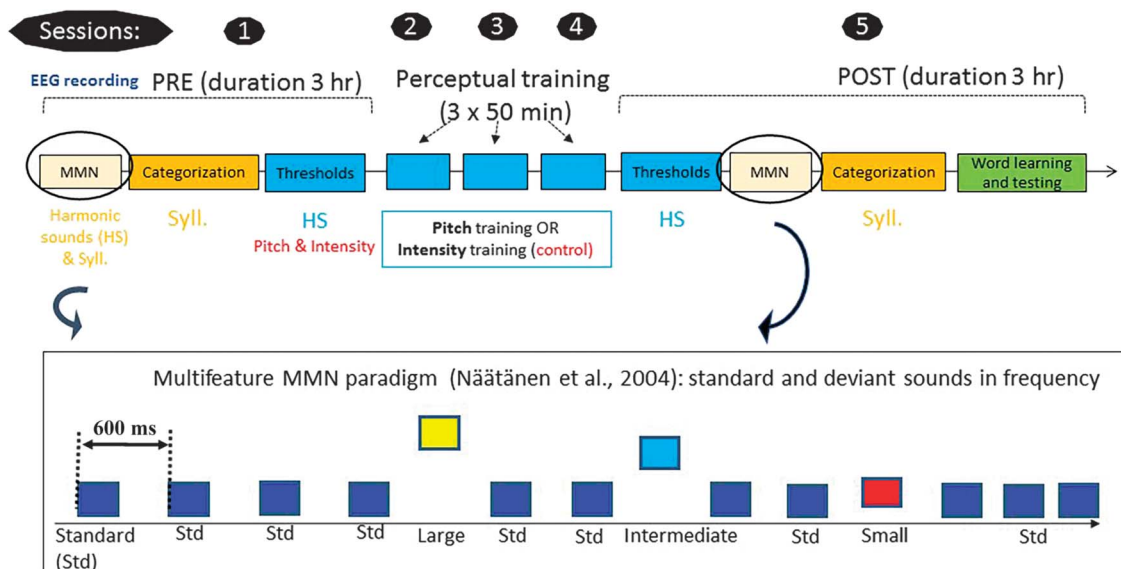
Most longitudinal studies typically focus on the impact of music training on the attentive processing of speech sounds. However, some longitudinal studies also examined the influence of music training on pre-attentive speech processing at the neural level using the brainstem auditory response (e.g., Tierney et al., 2013, 2015) or the mismatch negativity (MMN; Näätänen et al., 1978). Here, we used a longitudinal approach, and we analyzed the MMN to examine the influence of frequency training on pre-attentive sound processing. The MMN is a cortical correlate of automatic mismatch detection between rare deviant stimuli and more frequent standard stimuli in sound sequences. The MMN is computed as the difference wave of the event-related brain potentials (ERPs) elicited by the deviant minus the standard sounds. Importantly, the MMN is sensitive to training. For instance, early studies by Pantev and collaborators in adult nonmusicians showed that the MMN measured using magnetoencephalography (MEG, MMNm) was larger after 2 weeks of piano training compared with before training (Lappe et al., 2008, 2011). Using a longitudinal approach with children, Chobert et al. (2014) showed that the MMN, measured using electroencephalography (EEG) to syllables that differed in voice onset time (VOT) and syllabic duration, was larger in children who followed music training for 2 years but not in children involved in painting training. In the present experiment, we used the (EEG) MMN to examine transfer effects from frequency training with harmonic sounds to the pre-attentive processing of pitch variations in both harmonic sounds (near transfer) and syllables (far transfer).

This study is part of a larger project that aimed at better understanding the positive impact of musical

expertise on novel word learning. Previous cross-sectional studies have demonstrated that children, young adults, and to lesser extent old adults, who have been trained in playing a musical instrument for several years, learned the meaning of novel words from a tone language, Thai, in which pitch variations are linguistically relevant because they change the meaning of words, more efficiently than control participants without musical expertise (Dittinger et al., 2016, 2017, 2019). At the electrophysiological level, this facilitation was reflected by faster development of the N200 and N400 components of the ERPs in the word learning phase and faster integration of novel words' meaning into semantic networks in the test phase, in musicians than in nonmusicians. To determine whether these between-groups differences reflected the fact that greater sensitivity to pitch in musicians (Micheyl et al., 2006; Tervaniemi et al., 2005) causally drove successful learning of novel Thai words, we conducted a longitudinal study with two groups of nonmusician adult participants. One group of nonmusicians was trained on frequency, and the control group was trained on intensity using psychophysics procedures. In the pretraining session (D1, see Figure 1), the experimental design included a phonological categorization task, pitch and intensity threshold detection tasks, and an MMN experiment. Training included three sessions separated by at least 1 day (D2, D3, and D4) with participants trained either on frequency or on intensity for a total of 2.5 hr. Finally, in the posttest session (D5), participants performed again the phonological categorization task, the pitch and intensity threshold detection tasks, and the MMN experiment to analyze post- versus pretraining differences in the various measures of interest. Following these tasks, participants were involved in a novel word learning phase and in matching and semantic tasks to determine whether frequency training positively influenced semantic processing (see Figure 1). These results are reported in Barbaroux et al. (2021).

Here, we focus on the pre- and posttraining MMN experiments, performed on D1 and on D5, respectively. Each MMN experiment comprised two blocks, one that included standard and frequency deviant harmonic sounds and one that included standard and frequency deviant syllables. The main goal was to determine how pre-attentive processing of these two types of sounds is influenced by short-term psychoacoustic training (total duration around 2.5 hr) in frequency (experimental group) or in intensity (control group). Thus, we tested for transfer effects from frequency training to the perception of both harmonic sounds varying in frequency (near transfer) and syllables varying in frequency (far transfer). As is typical in MMN experiments using a multifeature design (Näätänen et al., 2004), deviant stimuli included large, intermediate, and small changes in frequency compared with the standard stimuli. This experimental procedure, with the presentation

Figure 1. Experimental design. Participants performed five sessions within 2 weeks. Of interest here, the mismatch negativity (MMN) experiment was performed during the first (preperceptual training) and last sessions (postperceptual training) using a multifeature MMN paradigm with standard harmonic sounds (HS) or syllables (/ba/) and three types of frequency deviant HS or syllables, which are large, intermediate and small, compared with standard sounds (see text for details). EEG = electroencephalography.



of frequent standard and rare deviant sounds, is known to generate not only MMN but also P3a component, as an involuntary orientation of attention toward novel and rare sounds (Donchin, 1981; Escera et al., 2000).

On the basis of previous results, we hypothesized that the MMN and the P3a component would be largest in amplitude and shortest in latency to large deviant stimuli, of intermediate amplitude and latency to intermediate deviant stimuli, and of smallest amplitude and longest latency to small deviant stimuli (main effect of deviance size; Tervaniemi et al., 2005). Moreover, we hypothesized that MMN and P3a amplitude would be larger post- than pretraining (main effect pre vs. post; Lappe et al., 2008, 2011). Most importantly, we predicted that the pre-attentive processing of the frequency deviant stimuli would be enhanced posttraining in the frequency training group but not in the intensity training group. We expected this training effect to be small because of the short duration of frequency training (total of 2.5 hr) but significant for large and intermediate pitch deviant stimuli. We did not expect this effect to be significant for small deviant stimuli because frequency training would be too short to improve the automatic detection of small frequency differences between the standard and deviant stimuli. In other words, we predicted a Training Group \times Pre-Post Interaction \times Deviant size interaction. Moreover, on the basis of the typical fronto-central distribution of the MMN and P3a component (Escera et al., 2000; Näätänen et al., 2004), we expected these interactive effects to be larger frontally than parietally (Training

Group \times Pre-Post \times Deviant Size \times Anterior-Posterior Interaction). Finally, if frequency training is causally related to both near and far transfer effects, the hypotheses above should be verified both for harmonic sounds (near transfer) and for syllables (far transfer). However, it is also possible that results only demonstrate near transfer effects or no transfer effects at all.

Method

Participants

A total of 28 nonmusician adults, all right-handed, participated in the study. Participants were monolingual but all had learned the basics of at least two foreign languages at school. They were pseudo-randomly assigned to one of two auditory training groups based on age, gender, and level of education, with 14 participants in the frequency group (FG; $M_{\text{age}} = 23.4$ years old) and 14 participants in the intensity group (IG; $M_{\text{age}} = 23.8$ years old, main effect of Group: $p = .84$) and eight women and six men in each group. None of the participants had prior experience with psychoacoustic tasks or with musical practice, except as part of their cursus in primary school. All participants had normal hearing, as defined by audiometric pure tone absolute thresholds below 20 dB HL at octave frequencies between 500 and 8000 Hz. The study was conducted in accordance with the ethical guidelines of Aix-Marseille University and

with the Helsinki declaration. All participants were paid for their participation. They signed an informed consent document, and they were told that they could stop the experiment at any moment if they felt uncomfortable.

Stimuli

Harmonic complex sounds. The standard stimulus had a f_o of 195 Hz and a duration of 326 ms. Three pitch deviant stimuli were created by increasing the f_o (using the Praat software; Boersma & Weenink, 2001) to 235 Hz for large deviant stimuli (+40 Hz, 20% increase), to 215 Hz for intermediate deviant stimuli (+20 Hz, 10% increase) and to 200 Hz for small deviant stimuli (+5 Hz, 2% increase). Harmonic sounds were played at an intensity level of 60 dB.

Syllables. The standard stimulus /ba/ had a f_o of 195 Hz, a vowel duration of 326 ms, and a VOT of -74 ms (total duration of 400 ms). Similar to harmonic sounds, three pitch deviants were created by increasing the f_o (using Praat) to 235 Hz for large deviants (+40 Hz, 20% increase), to 215 Hz for intermediate deviants (+20 Hz, 10% increase), and to 200 Hz for small deviants (+5 Hz, 2% increase), while leaving VOT and vowel duration with the same values as for the standard stimulus. Syllables were played at an intensity of 60 dB.

Procedure. For pre- and posttraining sessions, see Figure 1. Participants were involved in a longitudinal pretraining–training–posttraining procedure. In the pretraining session, participants performed several psychometric tests, phonological categorization tasks, psychoacoustic tests to measure pitch and intensity thresholds, and the MMN experiment. Half of the participants (FG) were trained on pitch discrimination, and the other half were trained on intensity discrimination (IG) within three psychoacoustic sessions over 2 weeks, on Days 2, 3, and 4 (three sessions at least 1 day apart), each lasting for 50 min (total training time 2.5 hr). The posttraining session (on Day 5) included the same tests as in the pretraining session together with novel word learning, matching, and semantic tasks. Here, we focus on the measure of the acoustic thresholds to ensure that psychoacoustic training showed the expected effects and on the MMN experiments to test for pre- versus posttraining differences.

Psychoacoustics. The psychoacoustical procedure to measure acoustic thresholds based on the Just Noticeable Difference (JND) method is described in detail in Barbaroux et al. (2021). Discrimination thresholds were measured using a two-interval, two-alternative forced-choice procedure with complex sounds. On each trial, two tones, the standard tone and the comparison tone, were presented in random order. In different blocks, tones varied in frequency (standard tone f_o 195 Hz + Δf Hz) or in intensity (standard tone intensity 60 dB + Δi dB). If the

first tone was higher than the second tone, participants in the FG were asked to press the left key, and if the second tone was higher than the first tone, they were asked to press the right key. Instructions were similar in the IG (if first tone was louder than the second tone, press the left key; if the second tone was louder than the first tone, press the right key). Pitch and intensity thresholds were measured in both the frequency and intensity training groups, and the order of the pitch and intensity discrimination tasks was counterbalanced across participants. A two-down, one-up procedure was used to estimate the frequency or intensity difference that corresponded to the 70% correct point on the psychometric function (Levitt, 1971).

MMN. Participants watched a self-selected silent movie displayed on a DVD-player screen, while complex harmonic sounds or syllables were presented binaurally through headphones (Sennheiser, HD600). They were told not to pay attention to the harmonic sounds or to the syllables that were presented in two different blocks lasting for 12.2 min each. A total of 1,200 stimuli were presented within each block, with 840 repetitions of the standard sound (including 15 consecutive standard sounds at the beginning of each block) and 120 repetitions (i.e., 10% of occurrence) of each deviant sound (large, intermediate, and small). Stimulus onset asynchrony was 600 ms.

EEG Data Acquisition

The EEG was continuously recorded at a sampling rate of 512 Hz from 32 active Ag-AgCl electrodes (Biosemi Pintype) located at standard position of the International 10/20 System (Jasper, 1958) using Biosemi amplifiers (Biosemi Active 2, Amsterdam, the Netherlands). Flat-type active electrodes were placed on the left and right mastoids as well as on the nose (reference electrodes). The electrooculogram was recorded from electrodes placed 1 cm to the left and right of the external canthi (to record horizontal ocular movements) and from an electrode beneath the left eye (for blinks). Electrode impedance was kept below 5 k Ω .

EEG data from nine electrodes (F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4) were analyzed using the Brain Vision Analyzer software (Version 1.05.0005; Brain Products, München, Germany). All data were rereferenced offline to the average of the left and right mastoids, filtered with a 0.1- to 40-Hz bandpass filter (12 dB/oct) and a notch filter at 50 Hz. Independent component analysis and inverse independent analysis were computed to remove components associated to horizontal and vertical eye movements. Recordings were segmented into 600 ms epochs (including a 100-ms baseline), time locked to stimulus onset. Direct current detrend and baseline corrections were applied to the segmented EEG signal, in addition to automatic removal of epochs containing artifacts (electrical activity

exceeding $\pm 75 \mu\text{V}$ around the baseline). Epochs were averaged within each condition to obtain individual averages and then averaged together across participants to obtain the grand average. Finally, the MMN and P3a components were computed for each participant by subtracting the ERPs to the standard stimulus from the ERPs to the deviant stimuli and then averaged together across participants to obtain mean difference waves.

Statistical Analyses

Because of a high level of noise in the EEG, four participants (two in each group) were excluded from the analyses, leaving 12 participants in each group. For both harmonic sounds and syllables, we analyzed the amplitude and latency of the MMN and P3a components using analysis of variance (ANOVA), computed with JASP (version 0.14.1; JASP Team, 2020). For harmonic sounds, the ANOVA included group (FG or IG) as a between-subjects factor and session (pre- vs. posttest), deviant stimuli (large, intermediate, or small), laterality (left, midline, and right) and anterior–posterior dimension (frontal, central, and parietal) as within-subject factors. For syllables, the MMN and the P3a components were only observed in response to large deviant stimuli. For the intermediate and small deviant stimuli, the MMN was not significantly different from zero. Consequently, ANOVA was only performed for large deviant stimuli, and they included group (FG or IG) as a between-subjects factor and session (pre- vs. posttest), laterality (left, midline, and right), and anterior–posterior dimension (frontal, central, and parietal) as within-subject factors. For each component of interest (i.e., MMN and P3a) and for each participant, the latency and maximum of amplitude (peak amplitude) were automatically detected in a latency range previously chosen based on averaged traces and visually checked for each participant (MMN to complex sound: 130–230 ms and MMN to syllables: 190–220 ms; P3a to complex sound: 250–350 ms and P3a to syllables: 290–330 ms). For each component, the mean amplitude was computed in a 20-ms latency band around the peak.

Results

Behavioral data. These results were presented in the study of (Barbaroux et al., 2021).

Pitch discrimination task. Overall, the between-groups difference was not significant (main effect of group: $F < 1$), but pitch thresholds decreased from pretraining (2.44 Hz) to posttraining (0.73 Hz, Tukey HSD, $p = .001$) in the pitch group, with no significant effect in

the IG (pre: 2.37 Hz; post: 1.84 Hz; Tukey HSD: $p = .55$; Group \times Session interaction, $F(1,24) = 4.57$, $p = .04$).

Intensity discrimination task. Overall, the between-groups difference was not significant (main effect of group: $F < 1$), but intensity thresholds decreased from pre-session (1.34 dB) to postsession (0.98 dB; $p = .001$) in the IG, with no significant effect in the pitch group (pre: 1.23 dB; post: 1.17 dB; $p = .55$; Group \times Session interaction: $F(1,24) = 6.52$, $p = .02$).

Electrophysiological MMN and P3a Component

Harmonic Sounds

Results for harmonic sounds in the frequency and intensity training groups are illustrated on Figures 2 and 3, respectively.

MMN Mean Amplitude

The MMN was larger frontally ($-2.27 \mu\text{V}$) than centrally ($-1.83 \mu\text{V}$) than parietally ($-1.09 \mu\text{V}$; main effect anterior–posterior: $F(2,44) = 96.54$, $p < .001$, $\eta^2 = .11$ post hoc tests, all $p < .001$). MMN was also larger to large ($-2.65 \mu\text{V}$) than to intermediate ($-1.32 \mu\text{V}$) than to small deviants ($-1.21 \mu\text{V}$; main effect of deviant: $F(2,44) = 54.91$, $p < .001$, $\eta^2 = .21$).

As expected based on our specific hypothesis, MMN was larger posttraining ($-1.92 \mu\text{V}$) than pretraining ($-0.53 \mu\text{V}$), main effect pre–post: $F(1,22) = 4.34$, $p < .049$, $\eta^2 = .02$. However, results revealed no between-groups difference (main effect of group: FG = $-1.64 \mu\text{V}$ and IG = $-1.82 \mu\text{V}$, $F < 1$) and no Group \times Pre–Post interaction ($F < 1$). The Group \times Pre–Post \times Anterior–Posterior interaction was significant, $F(2, 44) = 4.17$, $p < .022$, $\eta^2 = .001$, but post hoc tests did not reveal any effect of interest.

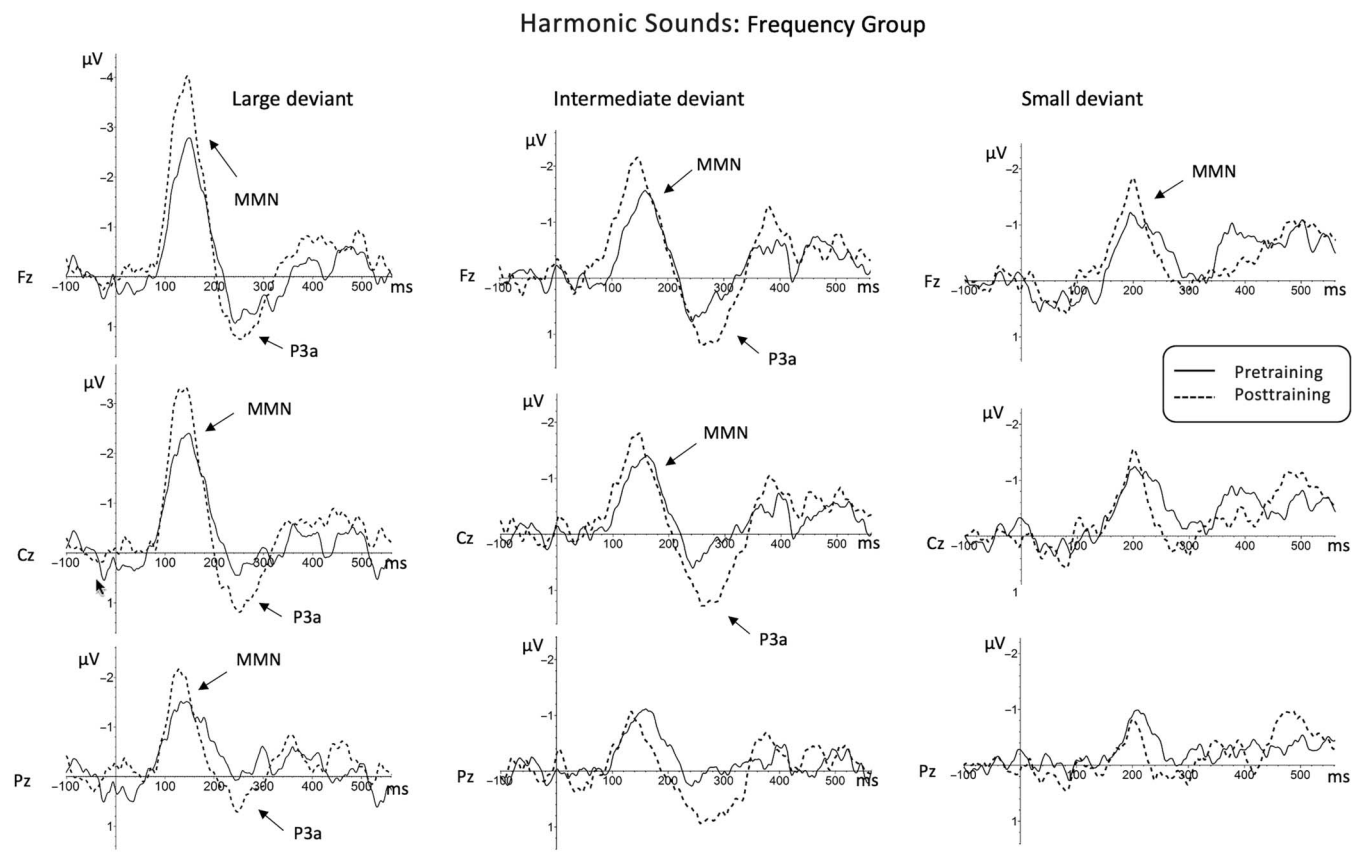
Preplanned pairwise comparisons were conducted to compare mean MMN amplitude pre- versus posttraining for each group (FG and IG) and for large and intermediate deviant stimuli. The MMN to large deviant harmonic sounds was larger posttraining than pretraining after frequency training only ($t = -2.37$, $p < .021$, $d = -0.84$). This difference was not significant after intensity training. Preplanned pairwise comparisons for intermediate deviants did not reach significance.

MMN Peak Latency

The main effect of anterior–posterior was not significant ($F < 1$). However, and as expected, MMN latency was shorter to large (140 ms) than to intermediate (158 ms) than to small deviants (201 ms; main effect of deviant = $F(2,44) = 174.08$, $p < .001$, $\eta^2 = .57$; post hoc tests: all $p < .001$).

As expected based on our specific hypothesis, MMN peak latency was shorter posttraining (163 ms) than

Figure 2. Difference waves for harmonic sounds in the frequency training group. Difference waves computed at frontal (Fz), central (Cz), and parietal sites (Pz) are overlapped preperceptual training (light black lines) and postperceptual training (light red lines) for large, intermediate, and small frequency deviant sounds. In this and subsequent figures, time in milliseconds is in abscissa and the amplitude of the effects in microvolt is in ordinate. Time zero corresponds to harmonic sound and syllables onset, and negativity is plotted upward. MMN = mismatch negativity.



pretraining (169 ms); main effect pre–post: $F(1,22) = 6.16$, $p < .021$, $\eta^2 = .008$. However, results revealed no between-groups difference (FG = 166 ms and IG = 167 ms; main effect of group: $F < 1$). The Group \times Pre–Post interaction was not significant, $F(1,22) = 2.56$, $p < .12$.

P3a Mean Amplitude

P3a was larger frontally (0.58 μV) and centrally (0.53 μV) than parietally (0.26 μV ; main effect anterior–posterior = $F(2,44) = 18.07$, $p < .001$, $\eta^2 = .021$, post hoc tests: both $p < .001$). P3a amplitude was also larger to large (0.81 μV) and to intermediate (0.57 μV) than to small deviants (–0.01 μV ; main effect of deviant = $F(2,44) = 16.23$, $p < .001$, $\eta^2 = .12$; post hoc tests, both $p < .001$).

As expected based on our specific hypothesis, P3a was larger posttraining (0.69 μV) than pretraining (0.23 μV); main effect pre–post: $F(1,22) = 16.78$, $p < .001$, $\eta^2 = .056$, across all scalp sites (Pre–Post \times Anterior–Posterior: $F < 1$). However, results revealed no between-groups difference (FG = 0.40 μV and IG = 0.52 μV , $F < 1$; main effect of group: $F < 1$). The Group \times Pre–Post interaction

and the Group \times Pre–Post \times Laterality interaction were not significant, $F(1,22) = 1.34$, $p = .26$ and $F(2, 44) = 1.60$, $p < .21$. The Group \times Pre–Post \times Anterior–Posterior interaction was significant, $F(2,44) = 3.30$, $p = .05$, $\eta^2 = .003$, but post hoc tests did not reveal any effects of interest.

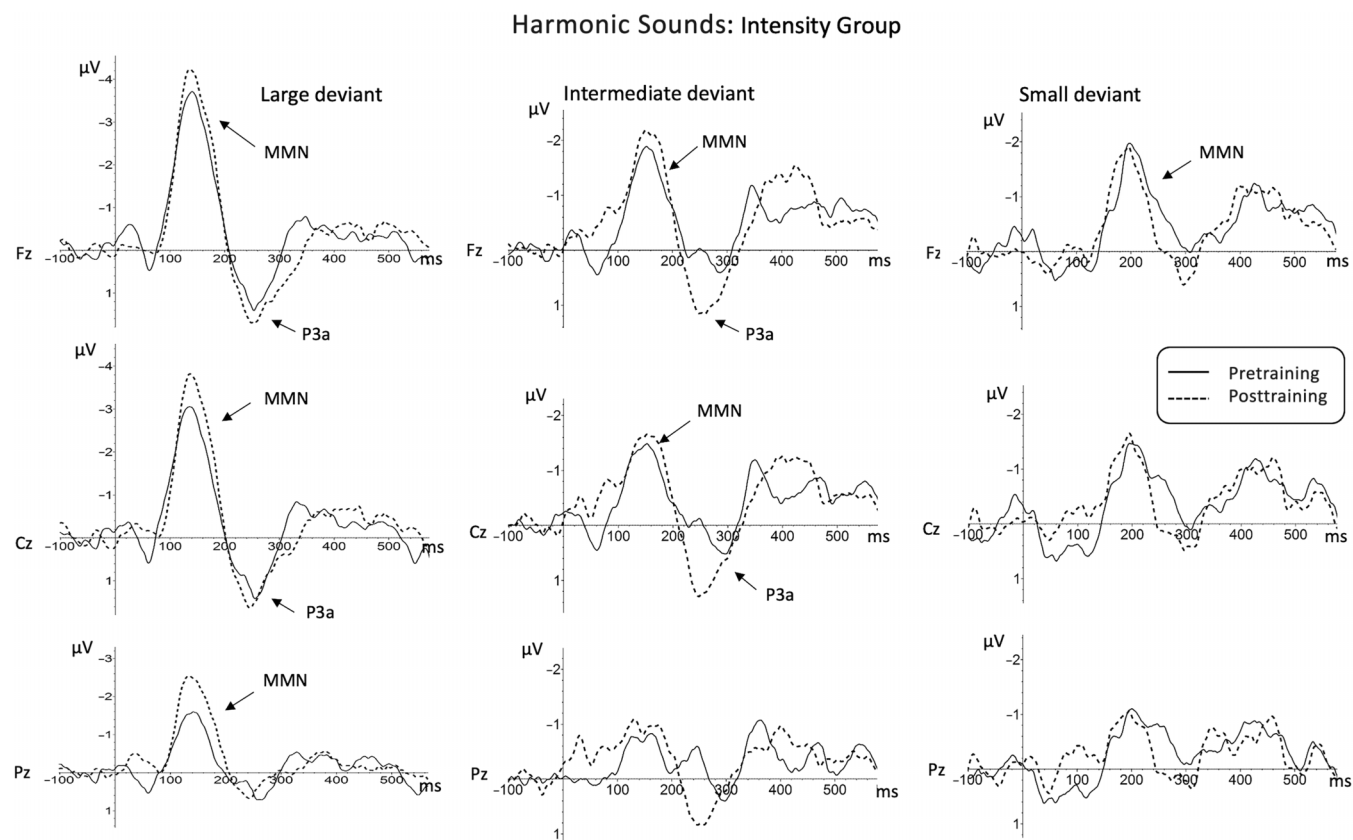
Finally, preplanned pairwise comparisons showed that P3a amplitude was larger posttraining than pretraining in the FG only for large deviants ($d = .63$, $t = 2.40$, $p < .019$) and in both groups for intermediate deviants (FG: $d = .91$, $t = 3.48$, $p < .001$, and IG: $d = .52$, $t = 1.98$, $p < .05$).

P3a Peak Latency

The main effect of anterior–posterior was not significant, $F(2,44) = 2.19$, $p < .124$. P3a latency was shorter to large (259 ms) than to intermediate (278 ms) than to small deviants (297 ms; main effect of deviant = $F(2,44) = 26.87$, $p < .001$; post hoc tests: all $p < .001$).

As expected based on our specific hypothesis, results showed that P3a latency was shorter posttraining (268 ms) than pretraining, 287 ms; main effect of pre–post:

Figure 3. Difference waves for harmonic sounds in the intensity training group. Difference waves computed at frontal (Fz), central (Cz), and parietal sites (Pz) are overlapped preperceptual training (light black lines) and postperceptual training (light red lines) for large, intermediate, and small frequency deviant sounds. MMN = mismatch negativity.



$F(1,22) = 30.99$, $p < .001$, $\eta^2 = .062$). However, results revealed no between-groups difference (GF = 277 ms and GI = 279 ms; main effect of group: $F < 1$). The Group \times Pre-Post interaction was not significant ($F < 1$).

Syllables

Results for syllables in the frequency and intensity training groups are illustrated on Figures 4 and 5, respectively. As no clear MMN was elicited by small or intermediate deviant syllables, analyses were only performed for large deviant syllables.

MMN Mean Amplitude

The MMN was larger frontally ($-1.51 \mu\text{V}$) and centrally ($-1.43 \mu\text{V}$) than parietally ($-0.94 \mu\text{V}$; main effect anterior-posterior: $F(2,44) = 19.68$, $p < .001$, $\eta^2 = .05$, post hoc tests, both frontal and central significantly larger than parietal ($p < .001$) and frontal not different from central, $p < .45$). Moreover, results of post hoc tests showed that only in the FG and in the pretesting session, the MMN was larger at frontal than at parietal electrodes

($p < .001$; Group \times Pre-Post \times Anterior-Posterior interaction: $F(2,44) = 4.76$, $p < .035$, $\eta^2 = .01$). The Group \times Pre-Post \times Laterality \times Anterior-Posterior interaction was also significant, $F(2,44) = 3.07$, $p < .035$, $\eta^2 = .001$, but none of the post hoc analyses revealed any significant results.

Results of preplanned pairwise comparisons were not significant.

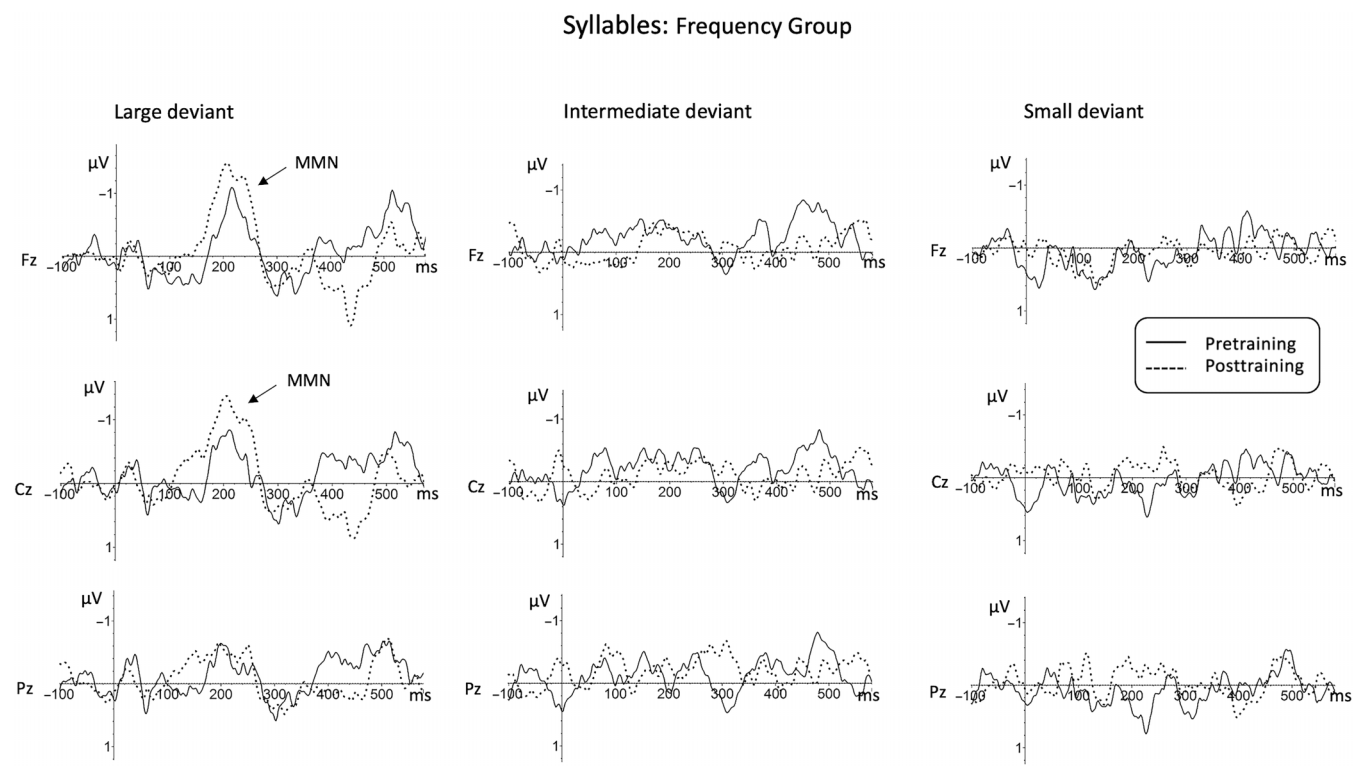
MMN Peak Latency

The main effect of anterior-posterior was significant, $F(2,44) = 5.39$, $p < .02$, $\eta^2 = .03$: The peak latency of the MMN was shorter parietally (202 ms) than frontally (213 ms; $t = 3.27$, $p < .006$). None of the other main effects or interactions reached significance.

P3a Mean Amplitude

Only the Pre-Post \times Laterality \times Anterior-posterior interaction, $F(4,88) = 3.44$, $p < .02$, $\eta^2 = .01$, was significant, but none of the post hoc tests reached significance. In contrast to our hypothesis, results of preplanned pairwise comparisons showed that P3a

Figure 4. Difference waves for syllables in the frequency training group. Difference waves computed at frontal (Fz), central (Cz), and parietal sites (Pz) are overlapped preperceptual training (light black lines) and postperceptual training (light red lines) for large, intermediate, and small frequency deviant syllables. MMN = mismatch negativity.



amplitude was larger posttraining than pretraining in the IG only ($d = .63$, $t = 2.43$, $p < .023$) and that the FG and IG were significantly different pretraining ($d = -.71$, $t = -2.01$, $p < .04$).

P3a Peak Latency

Results revealed no significant difference of interest.

Discussion

The aim of this experiment was to use a longitudinal test–training–retest procedure to determine whether short-term psychoacoustic training on the frequency or intensity attribute of harmonic sounds (in two different groups of non-musician participants) causally influenced the pre-attentive perception of harmonic sounds and of syllables varying in frequency. We recorded the brain electrical activity, and we compared the MMN and P3a component with large, intermediate, and small deviant stimuli pre- versus posttraining. In line with our hypothesis, results for harmonic sounds showed that the amplitude of both the MMN and P3a to large deviant stimuli were positively influenced by frequency training but not by intensity training, thereby evidencing near transfer effects. By contrast, no clear far

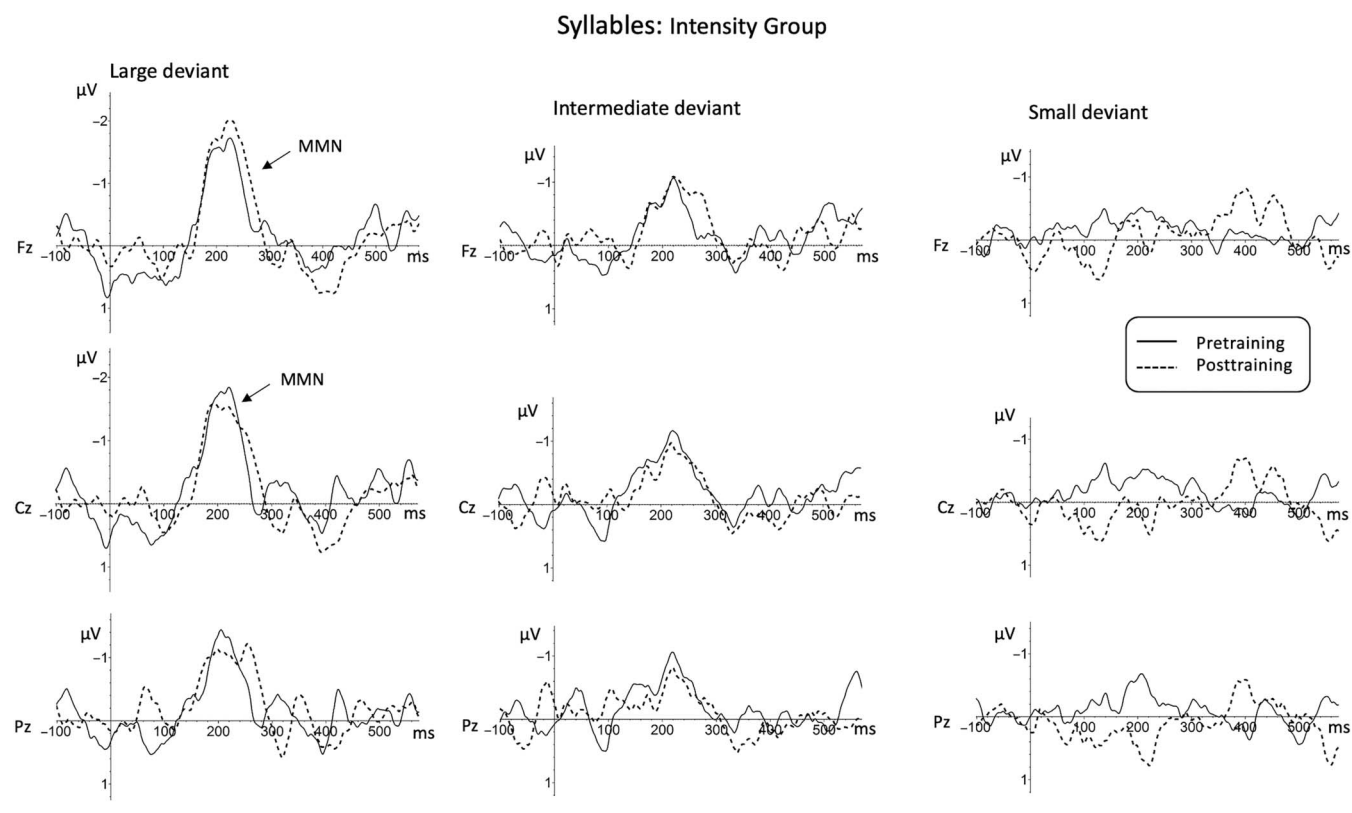
transfer effect was found for syllables, thereby possibly showing limits to transfer effects. These results are discussed below.

Harmonic Sounds

In line with previous results (Näätänen et al., 2004; Tervaniemi et al., 2005), the MMN and P3a measured in our experiments showed the typical frontal-larger-than-parietal scalp distribution. Moreover, in line with our specific hypothesis based on the study by Tervaniemi et al. (2005), both MMN and P3a amplitude and latency were larger/shorter to large than to intermediate than to small deviant sounds. Thus, these results demonstrated more efficient pre-attentive processing of harmonic sounds as a function of deviant size, with large changes in frequency between deviant and standard sounds (+40 Hz, 20% increase) generating stronger automatic MMN responses than intermediate (+20 Hz, 10% increase) and small (+5 Hz, 2% increase) frequency differences.

Most importantly, results for harmonic sounds revealed that both the MMN and P3a amplitude and latency were larger/shorter post than pretraining. In line with previous findings (e.g., Lappe et al., 2008, 2011), the present results showed that the MMN is a good index of

Figure 5. Difference waves for syllables in the intensity training group. Difference waves computed at frontal (Fz), central (Cz), and parietal sites (Pz) are overlapped preperceptual training (light black lines) and postperceptual training (light red lines) for large, intermediate, and small frequency deviant syllables. MMN = mismatch negativity.



brain plasticity. In the Lappe et al. (2008) experiment, nonmusician participants were trained either to play short musical sequences on the piano or to simply listen to these musical sequences to decide whether they were played correctly or not. In both groups, training was spread over 2 weeks with eight sessions of 25 min each, amounting to a total training time of around 3 hr, similar to the present experiment. The amplitude of the MMN recorded using MEG (MMNm) was enhanced posttraining compared with pretraining with larger MMN enhancement in the group that played compared with the group that only listened to the melodies, thereby showing that multimodal sensorimotor training induced stronger brain plasticity than auditory training only.

Evidence for Near Transfer Effects on the Pre-Attentive Processing of Large Frequency Deviant Stimuli

Results of preplanned comparisons confirmed one of our main hypotheses: The pre-attentive processing of large frequency deviant stimuli, as reflected both by MMN and P3a amplitude, was enhanced posttraining in the frequency training group but not in the intensity training group. This result is in line with the measures of pitch

thresholds showing a significant decrease from pretraining (2.44 Hz) to posttraining (0.73 Hz) in the FG but not in the IG (pre: 2.37 Hz and post: 1.84 Hz). Thus, a few hours of frequency training (around 2.5 hr over 2 weeks) allowed both to significantly decrease pitch discrimination thresholds in the FG and to enhance MMN and P3a amplitude to large frequency deviants (40 Hz, 20% increase).

However, in contrast to our hypothesis, neither the main effect of group nor the Group \times Pre-Post interaction was significant. Moreover, results for intermediate frequency deviants (20 Hz, 10% increase) showed no training effect on MMN amplitude, but P3a amplitude was larger posttraining than pretraining in both groups. Thus, while the duration of psychoacoustic training was possibly too short to specifically influence the automatic processing of relatively small changes in frequency, as reflected by the MMN, the general process of involuntary orientation of attention toward the deviant stimuli, as reflected by P3a amplitude, was enhanced by psychoacoustic training, independently of whether training was focused on frequency or on intensity (see Amitay et al., 2006, for similar results).

The finding that frequency training enhanced the MMN to large but not to intermediate deviant stimuli is surprising in view of the pitch thresholds reported above,

showing that after frequency training, participants detected frequency differences smaller than 1 Hz (0.73 Hz). Interestingly, Tervaniemi et al. (2005) found similar results in an experiment in which they presented a sequence of sounds that included standard stimuli and large, intermediate, and small frequency deviant stimuli. In one condition, participants, musicians and nonmusicians, were asked to detect the changes in frequency (attend condition). Results showed that musicians outperformed nonmusicians when the frequency deviant stimuli were difficult to detect. Nevertheless, and in line with our pitch threshold results, nonmusicians still reliably detected small 0.8% (4 Hz higher than the standard stimulus) and 2% (11 Hz higher than the standard stimulus) frequency changes between deviants and standard sounds. In the unattended condition, participants read a book while the same sequence of sounds was presented. In this condition, no significant difference was found between musicians and nonmusicians either in MMN or in P3a amplitude. Thus, while clear effects of music training in the Tervaniemi et al. (2005) experiment and of psychoacoustic training in our experiment were found when participants focused attention on the sound sequence and processed the sounds at the attentive level, training effects were not as strong at the pre-attentive level, when the sounds were automatically processed.

In sum, results for harmonic sounds showed that while psychoacoustic training was very efficient in decreasing pitch discrimination thresholds when the task required focused attention, the near transfer effect at the pre-attentive level of sound processing was only reliable for large changes in frequency. An obvious perspective for future research would be to determine whether stronger training effects at the pre-attentive level would develop with increased duration of psychoacoustic training.

Syllables

Because results showed no clear MMN and P3a component to intermediate and small deviant syllables, only the MMN and P3a to large deviant syllables were analyzed. In line with previous results, MMN amplitude was larger fronto-centrally than parietally (Escera et al., 2000; Näätänen et al., 2004; Tervaniemi et al., 2005), but MMN peak latency was shorter parietally than frontally. This unexpected result is possibly linked to the MMN not being clearly identifiable at parietal sites, at least in the frequency training group.

No Evidence for Far Transfer Effects on Pre-Attentive Syllabic Processing

Results revealed a significant Group \times Pre-Post \times Anterior-Posterior interaction. However, results of post hoc tests showed that the MMN was larger at frontal than at parietal electrodes in the FG in the pretesting session only but not posttraining as expected. Thus, in contrast to

our main hypothesis, results for syllables showed no significant effect of frequency training on MMN amplitude even for large frequency deviant syllables. Furthermore, results of preplanned comparisons showed that P3a amplitude was larger post- than pretraining in the IG only, but between-groups differences before training precluded to make strong inferences from these data.

General Discussion

In sum, results showed no far transfer effects from attentive psychoacoustic training on the frequency of harmonic sounds to the pre-attentive processing of the frequency of syllables. How can we account for this finding? Clearly, no strong conclusion can be drawn from a null result that implies accepting the null hypothesis (no between-groups differences). Nevertheless, because we found evidence for near transfer effects for large harmonic deviant sounds using the same procedure, it is of interest to speculate on the reasons why no far transfer effects were found for syllables.

First, the frequency differences between standard syllables and intermediate and small deviant syllables were clearly too small to generate reliable MMNs, and the amplitude of the MMN to large deviant syllables was also relatively small. For comparison purposes, we used the same deviance size values for syllables as for harmonic sounds. However, deviance size was obviously perceived differently at the pre-attentive level in both types of stimuli, possibly allowing for transfer of training effects to be found for large harmonic deviant sounds but not for syllables. This interpretation is in line with the observation that pitch variations are clearly more fine-grained in music than in speech (Patel, 2008; Schön et al., 2004). Thus, while we used the same frequency values with potentially different pitch perception thresholds for harmonic sounds and syllables, it would be of interest in future studies to examine training effects while equating pitch perception thresholds for both types of stimuli.

Second, the duration of frequency or intensity training (total 2.5 hr over 2 weeks) was possibly long enough to favor near transfer effects to large harmonic deviant sounds but too short to allow for far transfer effects to syllables. An obvious perspective for future research is to train participants for a longer duration to test for this interpretation.

Third, it is also possible that we found no far transfer effects because there is no shared processing of frequency in harmonic sounds and in syllables. For instance, previous results have shown that the level of performance in psychoacoustic tasks requiring frequency discrimination was unrelated to speech perception abilities (e.g., Boebinger et al., 2015; Surprenant & Watson, 2001). However, as

reviewed in the introduction, other results point to the positive influence of enhanced sensitivity to acoustic parameters in musicians on various aspects of speech perception (Delogu et al., 2010; Elmer et al., 2012; Lima & Castro, 2011; Wong et al., 2007). For instance, musicians, whether adults (Schön et al., 2004) or children (Magne et al., 2006), were more sensitive than nonmusicians to changes in the fundamental frequency of final sentence words, and this difference was reflected in the amplitude of their brain responses. Moreover, in a subsequent study, Moreno et al. (2009) used a longitudinal approach with nonmusician children to demonstrate that music training was indeed the cause of these differences. Nevertheless, it may be that music or frequency training has not the same impact when tested on low-level speech units, as syllables, and on large units such as words embedded in sentences (but see below).

Fourth, and as noted for harmonic sounds, transfer effects are possibly more difficult to evidence at the pre-attentive than at the attentive level of processing. This may be even more difficult for far than for near transfer effects. As mentioned above, this experiment was part of a larger project that included a phonological categorization task of Thai monosyllabic words based on pitch as well as a matching task and a semantic task. Results in these attentive tasks showed that participants were faster to categorize Thai monosyllabic words based on pitch after than before frequency training, with no significant differences after intensity training. Participants trained in frequency also made significantly fewer errors than participants trained in intensity when deciding that picture–word pairs matched or mismatched previously learned associations (Barbaroux et al., 2021). Therefore, in line with the bottom-up hypothesis described in the introduction, these results provided evidence that short-term psychoacoustic training with harmonic sounds positively impacted higher level of speech categorization and the learning of new associations when attention is focused on the task. Taken together, these results revealed fast transfer effects with short-duration low-level acoustic training when the task requires the attentive processing of speech sounds, but no transfer effects at the pre-attentive, automatic level of speech processing in the same time frame.

Limitations and Conclusions

Taken together, results of this experiment showed some evidence for near transfer effects of frequency training on the processing of large frequency deviant harmonic sounds but no evidence for far transfer effects on the processing of syllables, thereby showing some limits to transfer effects. As mentioned above, a first limitation is the short duration of psychoacoustic training that possibly impeded stronger near and far transfer effects. A second limitation rests on the choice to equate harmonic and syllabic deviant stimuli based on frequency rather than on pitch detection

thresholds. It will be of interest to overcome these limitations in future studies to better understand the various factors that influence transfer effects from low-level auditory abilities to higher levels of speech perception.

Author Contributions

Aline Frey: Formal analysis (Lead), Methodology (Lead), Validation (Supporting), Visualization (Lead), Writing – original draft (Supporting), Writing – review & editing (Supporting). **Mylène Barbaroux:** Conceptualization (Lead), Data curation (Lead). **Eva Dittinger:** Conceptualization (Supporting), Data curation (Supporting). **Mireille Besson:** Conceptualization (Lead), Data curation (Supporting), Formal analysis (Supporting), Funding, Methodology (Supporting), Supervision (Lead), Validation (Lead), Visualization (Supporting), Writing – original draft (Lead), Writing – review & editing (Lead).

Acknowledgments

The authors received funding from the Centre National de la Recherche Scientifique and Aix-Marseille University to conduct this research.

References

- Amitay, S., Irwin, A., & Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nature Neuroscience*, 9(11), 1446–1448. <https://doi.org/10.1038/nn1787>
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, 83(2), 111–130. [https://doi.org/10.1016/s0022-0965\(02\)00124-8](https://doi.org/10.1016/s0022-0965(02)00124-8)
- Barbaroux, M., Norena, A., Rasamimanana, M., Castet, E., & Besson, M. (2021). From psychoacoustics to brain waves: A longitudinal approach to novel word learning. *Journal of Cognitive Neuroscience*, 33(1), 8–27. https://doi.org/10.1162/jocn_a_01629
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Besson, M., Barbaroux, M., & Dittinger, E. (2017). Music in the brain: Music and language processing. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition* (pp. 37–48). Taylor & Francis Group. <https://doi.org/10.4324/9781315194738-4>
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, 2, 94. <https://doi.org/10.3389/fpsyg.2011.00094>

- Bidelman, G. M., & Alain, C.** (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *35*(3), 1240–1249. <https://doi.org/10.1523/JNEUROSCI.3292-14.2015>
- Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K.** (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378–387. <https://doi.org/10.1121/1.4904537>
- Boersma, P., & Weenink, D.** (2001). *Praat, a system for doing phonetics by computer* [computer program], version 4.0. <http://www.fon.hum.uva.nl/praat/>
- Chobert, J., François, C., Velay, J.-L., & Besson, M.** (2014). Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. <https://doi.org/10.1093/cercor/bhs377>
- Degé, F., & Schwarzer, G.** (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*, 124. <https://doi.org/10.3389/fpsyg.2011.00124>
- Delogu, F., Lampis, G., & Belardinelli, M. O.** (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, *22*(1), 46–61. <https://doi.org/10.1080/09541440802708136>
- Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M.** (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602. https://doi.org/10.1162/jocn_a_00997
- Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M.** (2017). Fast brain plasticity during word learning in musically-trained children. *Frontiers in Human Neuroscience*, *11*, 233. <https://doi.org/10.3389/fnhum.2017.00233>
- Dittinger, E., Korke, B., & Besson, M.** (2021). Evidence for enhanced long-term memory in professional musicians and its contribution to novel word learning. *Journal of Cognitive Neuroscience*, *33*(4), 662–682. https://doi.org/10.1162/jocn_a_01670
- Dittinger, E., Scherer, J., Jäncke, L., Besson, M., & Elmer, S.** (2019). Testing the influence of musical expertise on novel word learning across the lifespan using a cross-sectional approach in children, young adults and older adults. *Brain and Language*, *198*, 104678. <https://doi.org/10.1016/j.bandl.2019.104678>
- Donchin, E.** (1981). Surprise!... Surprise? *Psychophysiology*, *18*(5), 493–513. <https://doi.org/10.1111/j.1469-8986.1981.tb01815.x>
- Elmer, S., Meyer, M., & Jäncke, L.** (2012). Neurofunctional and behavioral correlates of phonetic and temporal categorization in musically trained and untrained subjects. *Cerebral Cortex*, *22*(3), 650–658. <https://doi.org/10.1093/cercor/bhr142>
- Escera, C., Alho, K., Schröger, E., & Winkler, I.** (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology & Neuro-Otology*, *5*(3–4), 151–166. <https://doi.org/10.1159/000013877>
- Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D.** (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLOS ONE*, *10*(9), e0138715. <https://doi.org/10.1371/journal.pone.0138715>
- Foncubierta, J. M., Machancoses, F. H., Buysse, K., & Fonseca-Mora, M. C.** (2020). The acoustic dimension of reading: Does musical aptitude affect silent reading fluency? *Frontiers in Neuroscience*, *14*, 399. <https://doi.org/10.3389/fnins.2020.00399>
- Frey, A., François, C., Chobert, J., Velay, J. L., Habib, M., & Besson, M.** (2019). Music training positively influences the preattentive perception of voice onset time in children with dyslexia: A longitudinal study. *Brain Sciences*, *9*(4), 91. <https://doi.org/10.3390/brainsci9040091>
- George, E. M., & Coch, D.** (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. <https://doi.org/10.1016/j.neuropsychologia.2011.02.001>
- Gomez-Dominguez, M., Fonseca-Mora, M. C., & Machancoses, F. H.** (2019). First and foreign language early reading abilities: The influence of musical perception. *Psychology of Music*, *47*(2), 213–224. <https://doi.org/10.1177/0305735617746734>
- Gordon, R. L., Fehd, H. M., & McCandliss, B. D.** (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*, 1777. <https://doi.org/10.3389/fpsyg.2015.01777>
- Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M.** (2016). Music and dyslexia: A new musical training method to improve reading and related disorders. *Frontiers in Psychology*, *7*, 26. <https://doi.org/10.3389/fpsyg.2016.00026>
- JASP Team.** (2020). *JASP* (Version 0.14.1)[Computer software]. <https://jasp-stats.org/>
- Jasper, H. H.** (1958). The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375. [https://doi.org/10.1016/0013-4694\(58\)90051-8](https://doi.org/10.1016/0013-4694(58)90051-8)
- Jentschke, S., & Koelsch, S.** (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, *47*(2), 735–744. <https://doi.org/10.1016/j.neuroimage.2009.04.090>
- Jeon, J. Y., & Fricke, F. R.** (1997). Duration of perceived and performed sounds. *Psychology of Music*, *25*(1), 70–83. <https://doi.org/10.1177/0305735697251006>
- Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C.** (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience*, *28*(39), 9632–9639. <https://doi.org/10.1523/JNEUROSCI.2254-08.2008>
- Lappe, C., Trainor, L. J., Herholz, S. C., & Pantev, C.** (2011). Cortical plasticity induced by short-term multimodal musical rhythm training. *PLOS ONE*, *6*(6), e21493. <https://doi.org/10.1371/journal.pone.0021493>
- Levitt, H.** (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2, Pt. 2), 467–477. <https://doi.org/10.1121/1.1912375>
- Lima, C. F., & Castro, S. L.** (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. <https://doi.org/10.1037/a0024521>
- Linnavalli, T., Putkinen, V., Lipsanen, J., Huotilainen, M., & Tervaniemi, M.** (2018). Music playschool enhances children's linguistic skills. *Scientific Reports*, *8*(1), 8767. <https://doi.org/10.1038/s41598-018-27126-5>
- Magne, C., Schön, D., & Besson, M.** (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. <https://doi.org/10.1162/089892906775783660>
- Marie, C., Magne, C., & Besson, M.** (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. <https://doi.org/10.1162/jocn.2010.21413>
- Michéyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J.** (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. <https://doi.org/10.1016/j.heares.2006.05.004>
- Moradzadeh, L., Blumenthal, G., & Wiseheart, M.** (2015). Musical training, bilingualism, and executive function: A closer

- look at task switching and dual-task performance. *Cognitive Science*, 39(5), 992–1020. <https://doi.org/10.1111/cogs.12183>
- Moreno, S., Lee, Y., Janus, M., & Bialystok, E.** (2015). Short-term second language and music training induces lasting functional brain changes in early childhood. *Child Development*, 86(2), 394–406. <https://doi.org/10.1111/cdev.12297>
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M.** (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19(3), 712–723. <https://doi.org/10.1093/cercor/bhn120>
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S.** (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R.** (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 115(1), 140–144. <https://doi.org/10.1016/j.clinph.2003.04.001>
- Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., Gabrieli, J., & Desimone, R.** (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences of the United States of America*, 115(28), E6630–E6639. <https://doi.org/10.1073/pnas.1808412115>
- Patel, A. D.** (2008). *Music, language, and the brain*. Oxford University Press.
- Patel, A. D.** (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, 2, 142. <https://doi.org/10.3389/fpsyg.2011.00142>
- Schellenberg, E. G.** (2004). Music lessons enhance IQ. *Psychological Science*, 15(8), 511–514. <https://doi.org/10.1111/j.0956-7976.2004.00711.x>
- Schön, D., Magne, C., & Besson, M.** (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341–349. <https://doi.org/10.1111/1469-8986.00172.x>
- Spiegel, M. F., & Watson, C. S.** (1984). Performance on frequency discrimination tasks by musicians and non-musicians. *The Journal of the Acoustical Society of America*, 76(6), 1690–1695. <https://doi.org/10.1121/1.391605>
- Straight, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R.** (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, 261(1–2), 22–29. <https://doi.org/10.1016/j.heares.2009.12.021>
- Surprenant, A. M., & Watson, C. S.** (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *The Journal of the Acoustical Society of America*, 110(4), 2085–2095. <https://doi.org/10.1121/1.1404973>
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E.** (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, 161(1), 1–10. <https://doi.org/10.1007/s00221-004-2044-5>
- Tierney, A. T., Krizman, J., & Kraus, N.** (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences of the United States of America*, 112(32), 10062–10067. <https://doi.org/10.1073/pnas.1505114112>
- Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N.** (2013). High school music classes enhance the neural processing of speech. *Frontiers in Psychology*, 4, 855. <https://doi.org/10.3389/fpsyg.2013.00855>
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N.** (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422. <https://doi.org/10.1038/nn1872>