Title	Statistical Learning of Chord-Transition Regularities in a Novel Equitempered Scale: An MMN Study
Author(s)	Ishida, Kai; Nittono, Hiroshi
Citation	Neuroscience Letters. 2023, 815, p. 137478
Version Type	AM
URL	https://hdl.handle.net/11094/92534
rights	© 2023. This manuscript version is made available under the CC-BY-NC-ND 4.0 license.
Note	

Statistical Learning of Chord-Transition Regularities in a Novel Equitempered Scale: An MMN Study

Short title: **STATISTICAL LEARNING OF CHORD-TRANSITION REGULARITIES**

Kai Ishida[1], Hiroshi Nittono[1]

[1] Graduate School of Human Sciences, Osaka University, Osaka, Japan

Corresponding author: Kai Ishida, Graduate School of Human Sciences, Osaka University, 1-2 Yamadaoka, Suita, Osaka 565-0871, JAPAN; ishida@hus.osaka-u.ac.jp; ORCID: 0000-0001-6485-0950

Hiroshi Nittono, Graduate School of Human Sciences, Osaka University, 1-2 Yamadaoka, Suita, Osaka 565-0871, JAPAN; nittono@hus.osaka-u.ac.jp; ORCID: 0000-0001-5671-609X

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding information: This work was supported by JSPS KAKENHI Grant Number JP22J20773.

Data accessibility:

The sound materials and dataset of the present paper are available at https://osf.io/p725n/.

CRediT author statement

Kai Ishida: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Project administration, Funding acquisition, Writing–Original Draft

Hiroshi Nittono: Conceptualization, Methodology, Writing–Review & Editing,

Highlights

- Chord sequences were created in a novel 18 equal temperament scale.
- Chords were presented with high or low transitional probabilities.
- Event-related potentials were recorded while participants listened to the chords.
- The chords with low transitional probability elicited mismatch negativity.
- Participants could not recognize the learned regularities beyond the chance level.

Keywords

statistical learning, mismatch negativity, event-related potential, musical chord sequence, Markov process

Abstract

In music and language domains, it has been suggested that patterned transitions of sounds can be acquired implicitly through statistical learning. Previous studies have investigated the statistical learning of auditory regularities by recording early neural responses to a sequence of tones presented at high or low transition probabilities. However, it remains unclear whether the statistical learning of musical chord transitions is reflected in endogenous, regularity-dependent components of the event-related potential (ERP). The present study aimed to record the mismatch negativity (MMN) elicited by chord transitions that deviated from newly learned transitional regularities. Chords were generated in a novel 18 equal temperament pitch class scale to avoid interference from the existing tonal representations of the 12 equal temperament pitch class system. Thirty-six adults without professional musical training listened to a sequence of randomly inverted chords in which certain chords were presented with high (standard) or low (deviant) transition probabilities. An irrelevant timbre change detection task was assigned to make them attend to the sequence during the ERP recording. After that, a familiarity test was administered in which the participants were asked to choose the more familiar chord sequence out of two successive sequences. The results showed that deviant transitions elicited the MMN, although the participants could not recognize the standard transition beyond the level of chance. These findings suggest that humans can statistically learn new transitional regularities of chords in a novel musical scale, even though they did not recognize them explicitly. This study provides further evidence that music-syntactic regularities can be acquired implicitly through statistical learning.

Introduction

In Western tonal music, each chord has a specific function (e.g., tonic, dominant and subdominant), and the arrangement of chords follows the rule of harmony. Several studies have shown that some aspects of tonal regularities in music can be acquired through long-term learning and plasticity [1–2]. For example, an event-related potential (ERP) study demonstrated that early right anterior negativity (ERAN), which reflects the violation of harmonic expectancy [3–4], shows a greater amplitude for rarer harmonic progressions based on the Western music corpus [5]. This result supports the notion of probabilistic or statistical learning of transitional regularities through everyday music listening.

Statistical learning is a process that realizes the grouping and segmentation of events in various sensory modalities based on probabilistic regularity and has been examined by behavioral and neural responses [6]. In the auditory domain, as evidence that transitional patterns of tones can be learned statistically, magnetoencephalography (MEG) studies have reported the attenuation of the magnetic counterparts of exogenous ERPs such as P1 and N1 for tones with higher transitional probability than for tones with lower transitional probability [7–8]. Moreover, recent studies have demonstrated that endogenous mismatch negativity (MMN) is elicited by tones with low transitional probability [9–12]. Because the MMN is the first endogenous component that reflects memory-based prediction based on the regularity extracted from the preceding context, the elicitation of the MMN can be a more direct indicator of statistical regularity learning than the attenuation of exogenous components, which are more dependent on sensory inputs.

Statistical learning has also been adopted for the acquisition of the regularity representation of note transition patterns within a melody and of chord-transition patterns [15–18]. Loui et al. [18] reported that infrequent chords in a novel musical scale (i.e., Bohlen-Scale) elicited early anterior negativity with a latency of 150–210 ms. They presented a chord progression in three different keys to enhance the generalizability of the progression pattern. However, it is still possible that the participants learned the patterns of pitch contours rather than the patterns of chord progressions.

In a musical context, a chord is defined as a simultaneously sounded harmonic set of tones with specific pitch classes rather than specific pitch heights. Even if two chords consist of tones with different pitch heights, they are categorized as the same chord if they consist of tones with the same pitch classes. For instance, the C major chord (C3–E3–G3) is still categorized as the C major chord, even if component C3 is raised by one octave (E3–G3–C4) and the chord is inverted. Considering this property of chords, Daikoku et al. [15] conducted an MEG study in which a particular chord was repeated three times in three different inversions and then transitioned to another chord with high or low probabilities. The effect of statistical learning was observed as the attenuation of an exogenous P1m for chords with high transitional probability compared to chords with low transitional probability around 70 ms after chord onset. However, Daikoku et al. did not examine the effect in a later latency range because they did not aim to record later components such as the MMN.

The present study aimed to investigate whether the statistical learning of the transitional regularity of chords is reflected in the preattentive MMN [13], the elicitation of which would provide additional evidence for statistical learning in addition to the attenuation of exogenous ERPs [15]. To avoid interference from existing tonal representations, an 18-equal temperament scale was used, and six types of triad chords were created. Triad chords were presented in three inversions to ensure that the target of learning was a harmonic chord (i.e., a set of pitch classes) rather than a set of tones with specific pitch heights [15]. To record the MMN, the experimental paradigm of Koelsch et al. [9] and Tsogli et al. [12] was adopted. In their studies, statistical learning of regularities, as indicated by the MMN, was implicit, because the participants were not aware of the regularities. In the present study, the transitional regularities of chords were manipulated. Various chord triplets were presented repeatedly without a pause. In each chord triplet, the first two chords formed a “root,” and the last chord was an “ending.” Each root transitioned to one of two types of endings at high ($p = .90$) or low ($p = .10$) probabilities so that the same ending became either a standard or deviant transition, depending on the roots. Because each chord was presented with equal probability and the same chords would become either the standard or deviant, the MMN observed in the present study could reflect deviant detection based on the regularity acquired through statistical learning, rather than the processing of the occurrence frequency of chords or the change in acoustic features. After the ERP experiment, the implicit or explicit nature of learning was examined using a familiarity test to determine whether participants could recognize the transitional regularities above the level of chance.

Materials and Methods

Participants

The sample size ($N = 34$) was determined using G*Power [19] to detect a medium effect size ($dz = 0.5$) with a power of .80. This medium effect size was selected according to the effect sizes reported in previous studies, which were often larger than 0.5 ($dz = 0.76$ [9] or 1.33 [12] for MMN; $dz = 0.69$ [18] for the early anterior negativity). Expecting some dropouts, we recruited 36 adults without professional musical training (23 women and 13 men, 18–38 years old, $M = 22.6$ years). However, all participants' data could be used for hypothesis testing. All participants were right-handed [20], and none had hearing impairments or a history of neurological disease. The participants had various types of musical experience, with a mean of 5.8 years of extracurricular musical lessons (range 0–17 years). The protocol was approved by the Behavioral Research Ethics Committee of the Osaka University School of Human Sciences, Japan (HB022-107), and written informed consent was obtained from all participants. Participants received a cash voucher of 3,500 Japanese yen as an honorarium.

Figure 1

Materials

Figure 1 shows all chords used in the present study. Six types of triad chords consisting of notes from the 18 equal-temperament scale were created in three different inversions while controlling interference by the Western music corpus [15]. The rationale for using this scale was described in Supplementary Material. Each note was a sine tone to avoid a timbre-specific effect. In the chord sequence, the duration of a chord was 450 ms, which included a rise and fall of 10 and 200 ms. Each chord was presented with an interstimulus interval of 50 ms (thus, the inter-onset interval was 500 ms). All chords were sampled at 44,100 Hz, and the amplitude was normalized.

All methods were similar to those in Tsogli et al. [12]. Three chords were connected to form a triplet. Four types of triplets were created by combinations of two types of roots (AC or BD) and two types of endings (E or F). Each root transitioned to either E or F chords with high or low frequency. The left panel of Figure 2 shows how the chords transitioned. To control for the combination effect of chords, whether the chords were high or low transitions was counterbalanced between participants. In Group I, the AC transitioned to the E chord with high probability ($p = .90$) and transitioned to the F chord with low probability ($p = .10$), while the BD transitioned to the F chord with high probability and transitioned to the E chord with low probability. In Group II, the AC transitioned to the F chord with high probability and to the E chord with low probability, while the BD transitioned to the E chord with high probability and to the F chord with low probability. The transitional probability between the triplets was equal ($p = .50$). Thus, all chords had an equal probability of occurrence. Note that three types of inversions of a chord were randomly presented at each chord position.

Procedure

In the EEG recording, all four triplets were presented in random order.[1] During the presentation, the participants were asked to press a button as quickly and accurately as possible when the chords changed to a piano in a timbre change detection task. Using this kind of detection task is a common method in statistical learning to make participants pay attention to a stimulus sequence without focusing on its regularities [7–9, 12, 15]. Timbre changes occurred at random positions within a triplet 3–5 times in each block for a total of 40 changes in the entire experiment. The timbre change occurred only in triplets including chords with high transitional probability to avoid reducing the number of trials with low transitional probability. Ten blocks were performed with short breaks, and each block lasted approximately six minutes. Within a block, triplets whose roots were AC and BD were randomly presented 1,000 times each,

---

[1] After this session, we conducted another block in which each triplet was presented with equal probability (about 20 minutes). However, the data will not be reported here because we failed to randomize chord inversions and the results were uninterpretable.

with the constraint that triplets with low transitional probabilities were not repeated in succession. In Group I, ACE, ACF, BDF, and BDE were presented 900, 100, 900, and 100 times, respectively. In Group II, ACF, ACE, BDE, and BDF were presented 100, 900, 100, and 900 times, respectively.

The EEG recording was followed by a two-alternative forced-choice familiarity test that took four minutes. In the familiarity test, four types of possible pairs of the unlearned triplets (ACE vs. BDE, BDE vs. ACE, ACF vs. BDF, and BDF vs. ACF) were presented six times (i.e., 24 trials in total). The order of presentation of the roots was counterbalanced across participants. The pause between two triplets of a pair was 500 ms. The participants' task was to choose which triplets sounded more familiar by pressing a key that corresponded to either the first or second triplet. The choice of the triplet that contained chords with high transitional probability was regarded as the correct response. After choosing the triplet, participants described their confidence in their choice using a scale from 1 = *very unsure* to 5 = *very sure* at their own pace. The regularity of chord transitions was explained at the end of the experiment.

EEG recording and data reduction

EEG data were recorded using a QuickAmp (Brain Products) with Ag/AgCl electrodes. Thirty-four scalp electrodes were applied according to the 10–20 system (Fp1/2, F3/4, F7/8, Fz, FC1/2, FC5/6, FT9/10, C3/4, T7/8, Cz, CP1/2, CP5/6, TP9/10, P3/4, P7/8, Pz, O1/2, Oz, PO9/10). Additional electrodes were placed on the left and right mastoids, the left and right outer canthi of the eyes, and above and below the right eye. The data were referenced offline to the algebraic means of the left and right mastoid electrodes. The sampling rate was 1,000 Hz. The online filter was DC–200 Hz. Electrode impedances were kept below 10 kΩ.

EEG data were analyzed using Brain Vision Analyzer (Brain Products, Germany). First, a digital filter of 0.5 Hz (6 dB/oct) high-pass filter and 30 Hz (48 dB/oct) low-pass filter and a notch filter of 60 Hz were applied to the data [9, 12]. After correcting ocular and other artifacts (see Supplementary Material for details), a 500 ms period (100 ms before and 400 ms after the ending note) was averaged after removing trials in which voltages exceeded ±80 μV in any channel. Two consecutive trials after the timbre change were removed from the analysis. Baseline correction was applied by subtracting the mean amplitude of the prestimulus 100 ms from each point of the waveform. Statistical evaluation was conducted at the frontal electrode cluster (F7, F3, Fz, F4, and F8) based on the previous study [21]. The peak of MMN was detected in the interval of 150–280 ms of the grand mean difference waveform (using all 10 blocks), calculated by subtracting the ERP waveform of chords with high transitional probability from that of chords with low transitional probability. The interval ± 20 ms from the peak was defined as the MMN interval. On average, 1644 (1496–1670) and 197 (177–200) epochs were used to calculate the standard and deviant ERP waveforms, respectively.

Statistical analysis

Statistical analyses were carried out using JASP 0.17.2 [22]. A mixed two-way analysis of variance (ANOVA) with condition (standard vs. deviance), and group (Group I vs. Group II) was conducted on the ERP amplitude of the MMN interval. This analysis was also conducted using a Bayesian mixed two-way ANOVA to assess the absence (effect size $\delta = 0$, null hypothesis) or presence (effect size $\delta \neq 0$, alternative hypothesis) of the effects. The correct percent of the familiarity test was aggregated across the groups, and compared to the chance level ($p = .50$) using a one-sample t-test (one-sided) because the correct percentages of both groups were not significantly different, $t(31.5) = 0.114$, $p = .910$, Cohen's $d = .038$, $BF_{01} = 3.093$. The same analysis was conducted using a Bayesian one-sample t-test to assess the absence (effect size $\delta = 0$, null hypothesis) or presence of the difference (effect size $\delta < 0$, alternative hypothesis). Finally, the confidence rating between correct and incorrect responses was compared using a paired t-test (two-sided) and a Bayesian paired t-test. For frequentist hypothesis testing, the significance levels were set to $\alpha = .05$. For Bayesian hypothesis testing, the Cauchy distribution with a scale parameter r of 0.707 was used as the prior distribution for δ in the t-test. For the Bayesian two-way repeated-measures ANOVA, multivariate Cauchy distribution (fixed effect: scale parameter $r = 0.5$; random effect: scale parameter $r = 1$; covariates: scale parameter $r = .354$) was used as the prior distribution. As an exploratory analysis, MMN amplitudes in the former and latter halves of the experiment were compared to examine the effect of learning (see Supplementary Material).

Results

The averaged mean reaction time of the timbre change detection task was 308 ms (*SD* = 46 ms), and the averaged hit rate was 98.9% (*SD* = 3.0%), suggesting that the participants focused on the task and attended to the chord sequence.

Figure 2

The right panel of Figure 2 shows the grand average waveforms and scalp topographies of the ERPs. Chords with low transitional probability elicited the MMN ($M = -0.223$ μV, $SD = .487$) over the frontal area, with a peak latency of 206 ms. Therefore, a period of 186–226 ms was used for scoring MMN amplitudes. Similar to the typical MMN, a slight polarity inversion was observed when the data were referenced to the nose (see Supplementary Figure S2). The mixed two-way ANOVA conducted on the MMN amplitudes revealed the significance of condition, $F(1, 34) = 7.331$, $p = .011$, $\eta_p^2 = .177$, $BF_{10} = 4.532$, suggesting that the MMN was elicited by the deviant chord transition irrespective of the combination of the chord. None of the other effects and interactions were significant, $F(1, 34) < 0.368$, $p > .548$, $\eta_p^2 < .011$, $BF_{10} < 0.561$. Although the MMN was observed, the one-sample *t*-test showed that the percentage of correct responses ($M = 52.5\%$, $SD = 12.0$) did not significantly exceed the chance level, $t(35) = 1.277$, $p = .105$, Cohen's $d = 0.213$, $BF_{+0} = 0.671$. The difference between confidence ratings when the response was correct ($M = 3.3$, $SD = 0.6$) or incorrect ($M = 3.2$, $SD = 0.5$) was not significant, $t(35) = 0.741$, $p = .463$, Cohen's $d = 0.124$, $BF_{10} = 0.231$.

Discussion

The present study examined whether the MMN response is elicited by deviations from the statistically learned transitional regularity of chords, defined as a harmonic set of pitch classes in a novel musical scale. The results of the ERP showed that a chord elicited MMN when it was presented with a low transitional probability, even if the chord was presented equiprobably in the whole experiment. The results of the familiarity test, however, showed that the participants could not recognize the standard transition beyond the level of chance, and there was no difference in confidence ratings between correct and incorrect responses, suggesting that the participants chose the triplets without clear response criteria. These behavioral results indicate that the acquired representation was implicit.

This study provides further evidence that the transitional regularities of chords are statistically learned by demonstrating that the MMN, which is a memory-based endogenous component [14], is elicited by chords with low transitional probability. Consistent with the MMN around 180–260 ms after the onset of the tone of Tsogli et al. [12], the latency of the MMN response (186–226 ms) in the present study was later than that of the traditional MMN response, such as a physical MMN elicited by the infrequent change in the acoustic feature of tone (e.g., 100–200 ms; as reviewed by 13–14). Tsogli et al. suggested that the generation of the statistical MMN requires more top-down processing to encode the deviance than the physical MMN because a longer time is needed to learn the contextual regularity. This coincides with previous findings that the latency of the MMN is longer as the complexity of stimuli increases [23]. Moreover, the exploratory analysis showed that a significant MMN occurred only in the latter half of the experiment, which supports the learning effect (see Supplementary Material).

The MMN response of the present study and the early anterior negativity reported by Loui et al. [18] may be the same kind of component that reflects auditory deviant detection based on the statistical learning of chord progressions or transition regularities. By manipulating chord inversions, the present study extends the findings of Loui et al. to the more abstract harmonic regularity, where chords are defined as a harmonic set of pitch classes in a novel musical scale. Furthermore, the generation process of the MMN in the current study may be similar to that of the ERAN. The ERAN reflects the violation of harmonic expectancy based on the schema of musical syntax acquired as a long-term format [2]. MMN is thought to reflect the innate ability to extract regularities in the relationship between sounds and immediately establish representations, as MMN-like discriminative responses are elicited by changes in acoustic features from the infant stage [24]. Through a review of ERAN studies, Koelsch [2] noted that the formation and organization processes of auditory objects, which are required to generate the MMN, are indispensable for conducting music-syntactic processing. Furthermore, some studies have proposed that the ERAN and MMN reflect similar irregularity detection processes, despite differences in regularity representations [4, 21]. In the present study, the learning processes of chord transitions may be similar to the process required to acquire the pattern representations of harmony. Taken together, the regularity representations underlying music-syntactic processing are possibly acquired by statistical or probabilistic learning [2, 5].

In the present study, the participants could not recognize the regular transition pattern beyond the chance level. Previous studies have also reported that the MMN can be elicited when the performance on the follow-up behavioral test is below the chance level [9, 11–12]. ERPs have been considered a more sensitive measure of statistical learning than behavioral measures [25]. It is also possible that the discrepancy between neural and behavioral results may reflect insufficient learning. The MMN amplitude of the present study was smaller than that of the early anterior negativity of Loui et al. [18], in which explicit recognition beyond the chance level was achieved. The regularity in the present musical stimuli might be difficult to learn sufficiently in one hour of listening. Moreover, musical proficiency may facilitate statistical learning [8]. Because the present study did not control for participants' musical experience and absolute pitch ability, future research is needed to examine the relationship between ERP and behavioral measures in statistical learning of musical stimuli.

In summary, the present results demonstrated that chords with low transitional probability elicited the MMN. This is due to the statistical learning of transitional regularities of chords (defined as a harmonic set of pitch classes) in a novel musical scale. The participants could not recognize the standard transition chords beyond the level of chance. Future neuroscientific research should examine whether explicit knowledge of regularity can be acquired when the regularity is learned intentionally. In conclusion, the present study suggests that the representation of music-syntactic regularities can be acquired through statistical learning.

References

[1] H. Honing, C. ten Cate, I. Peretz, & S. E. Trehub, Without it no music: Cognition, biology and evolution of musicality. Philos. Trans. R. Soc. 370 (2015) 20140088. https://doi.org/10.1098/rstb.2014.0088

[2] S. Koelsch, Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. Psychophysiology 46 (2009) 179–190. https://doi.org/10.1111/j.1469-8986.2008.00752.x

[3] S. Koelsch, T. Gunter, A. D. Friederici, & E. Schröger, Brain indices of music processing: “Nonmusicians” are musical. J. Cogn. Neurosci. 12 (2000) 520–541. https://doi.org/10.1162/089892900562183

[4] S. Koelsch, T. C. Gunter, E. Schröger, M. Tervaniemi, D. Sammler, & A. D. Friederici, Differentiating ERAN and MMN: An ERP study. NeuroReport 12 (2001) 1385–1389. https://doi.org/10.1097/00001756-200105250-00019

[5] S. G. Kim, J. S. Kim, & C. K. Chung, The effect of conditional probability of chord progression on brain response: An MEG study. PLOS ONE 6 (2011) e17337. https://doi.org/10.1371/journal.pone.0017337

[6] C. M. Conway, How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. Neurosci Biobehav Rev. 112 (2020) 279–299.

[7] T. Daikoku, Y. Yatomi, & M. Yumoto, Implicit and explicit statistical learning of tone sequences across spectral shifts. Neuropsychologia 63 (2014) 194–204. https://doi.org/10.1016/j.neuropsychologia.2014.08.028

[8] T. Daikoku, & M. Yumoto, Musical expertise facilitates statistical learning of rhythm and the perceptive uncertainty: A cross-cultural study. Neuropsychologia,

146 (2020) 107553. https://doi.org/10.1016/j.neuropsychologia.2020.107553

[9] S. Koelsch, T. Busch, S. Jentschke, & M. Rohrmeier, Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. Sci. Rep. 6 (2016) 19741. https://doi.org/10.1038/srep19741

[10] T. Moldwin, O. Schwartz, & E. S. Sussman, Statistical learning of melodic patterns influences the brain's response to wrong notes. J. Cogn. Neurosci. 29 (2017) 2114–2122. https://doi.org/10.1162/jocn

[11] E. Paraskevopoulos, A. Kuchenbuch, S. C. Herholz, & C. Pantev, Statistical learning effects in musicians and non-musicians: An MEG study. Neuropsychologia 50 (2012) 341–349. https://doi.org/10.1016/j.neuropsychologia.2011.12.007

[12] V. Tsogli, S. Jentschke, T. Daikoku, & S. Koelsch, When the statistical MMN meets the physical MMN. Cogn. Sci. 9 (2019) 5563. https://doi.org/10.1038/s41598-019-42066-4

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

[14] E. S. Sussman, S. Chen, J. Sussman-Fort, & E. Dinces, The five myths of MMN: Redefining how to Use MMN in basic and clinical research. Brain Topogr. 27 (2014) 553–564. https://doi.org/10.1007/s10548-013-0326-6

[15] T. Daikoku, Y. Yatomi, & M. Yumoto, Pitch-class distribution modulates the statistical learning of atonal chord sequences. Brain Cogn. 108 (2016) 1–10. https://doi.org/10.1016/j.bandc.2016.06.008

[16] E. M. M. Jonaitis, & J. R. Saffran, Learning harmony: The role of serial statistics.

Cogn. Process. 33 (2009) 951–968. https://doi.org/10.1111/j.1551-6709.2009.01036.x

[17] P. Loui, Learning and liking of melody and harmony: Further studies in artificial grammar learning. Top. Cogn. Sci. 4 (2012) 554–567. https://doi.org/10.1111/j.1756-8765.2012.01208.x

[18] P. Loui, E. H. Wu, D. L. Wessel, & R. T. Knight, A generalized mechanism for perception of pitch patterns. J. Neurosci. 29 (2009) 454–459. https://doi.org/10.1523/JNEUROSCI.4503-08.2009

[19] F. Faul, E. Erdfelder, A. G. Lang, & A. Buchner, G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39 (2007) 175–191. https://doi.org/10.3758/bf03193146

[20] R. C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1971) 97-113. https://doi.org/10.1016/0028-3932(71)90067-4

[21] K. Ishida, & H. Nittono, Relationship between early neural responses to syntactic and acoustic irregularities in music. Eur. J. Neurosci. 56 (2022) 6201–6214. https://doi.org/10.1111/ejn.15856

[22] JASP Team, JASP (version 0.17.2) [computer software] (2023). https://jasp-stats.org/.

[23] T. Kujala, M. Tervaniemi, & E. Schröger, The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. Biol. Psychol1. 74 (2007) 1–19. https://doi.org/10.1016/j.biopsycho.2006.06.001

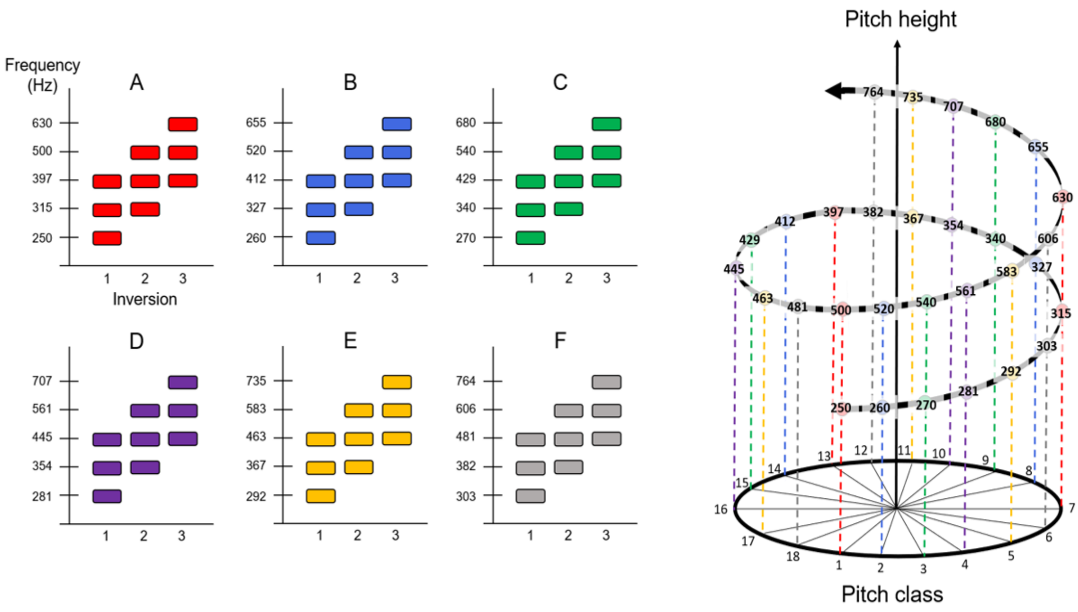
[24] M. Cheour, P. H.t. Leppänen, & N. Kraus, Mismatch negativity (MMN) as a tool

for investigating auditory discrimination and sensory memory in infants and children. Clin. Neurophysiol. 111 (2000) 4–16. https://doi.org/10.1016/S1388-2457(99)00191-1

[25] T. Okano, T. Daikoku, Y. Ugawa, K. Kanai, & M. Yumoto, Perceptual uncertainty modulates auditory statistical learning: A magnetoencephalography study. *Int J Psychophysiol.* 168 (2021) 65–71. https://doi.org/10.1016/j.ijpsycho.2021.08.002

Figure 1.

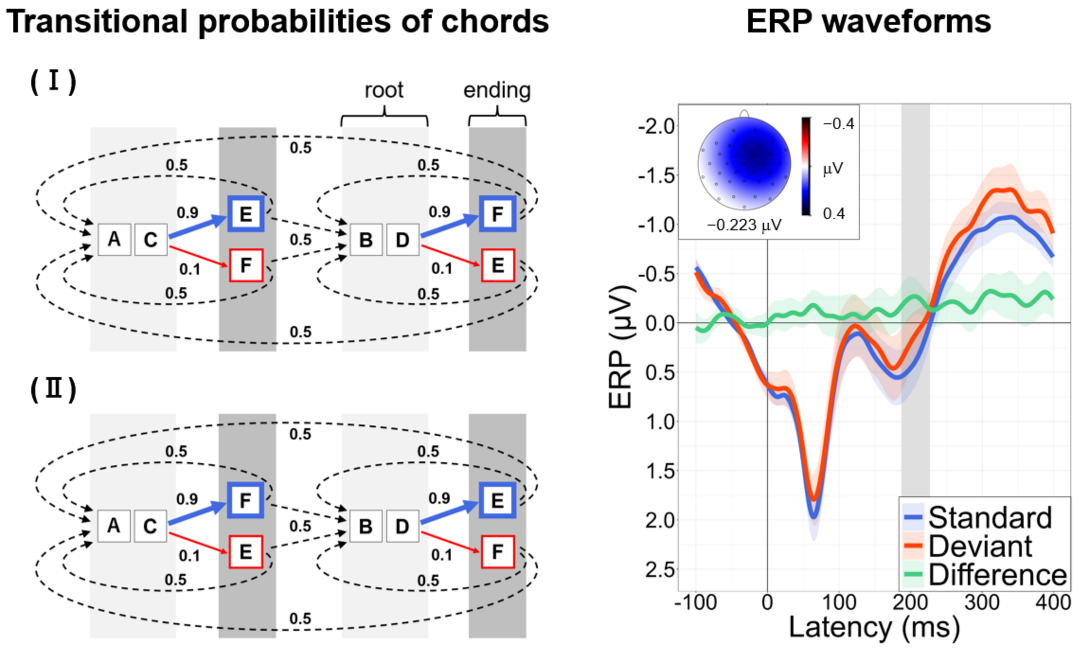
Chords used in the present study and the pitch helix of 18 equal temperament



Note. The left panel shows the six types of triad chords and their three inversions. The vertical axis indicates the frequency (Hz), and the horizontal axis indicates the versions of chord inversion. For example, chord A could be [250, 315, 397 Hz], [315, 397, 500 Hz], or [397, 500, 630 Hz]. The right panel shows the pitch helix of 18 equal temperaments. Each dot indicates the pitch used to construct each chord, and the numbers indicate the frequency of each pitch.

Figure 2.

Transitional probabilities of the chords and the grand average ERP waveforms


Transitional probabilities of chords

ERP waveforms

Note. In the left panel, Roman numerals in parentheses of the top and bottom figures indicate the group. In the right panel, grand average waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) with 95% confidence intervals and the topographic map (186–226 ms) of the deviant-related difference waveforms are shown.

Supplementary Material

Rationale for using the 18 equal temperament scale

In Koeslch et al. [9] and Tsogli et al. [12], six different timbres transitioned with high or low probabilities. To reproduce this type of regularity in chords defined as a set of pitch classes, we created six types of triad chords each consisting of three pitch classes. The 18 equitempred scale was required to define each triad chord as a distinct set of unique pitch classes (3 × 6 pitch classes). Each of the six chords was presented randomly in three different inversions to avoid pitch-specific learning.

Artifact Correction

EEG data were preprocessed using the *Ocular Correction ICA* (independent component analysis) function of Brain Vision Analyzer 2.2 (Brain Products, Germany). The InfoMax algorithm was used. The dataset of 41 channels (i.e., 34 scalp, four EOG, two mastoid, and one nose channels) was analyzed. Detection of ICs associated with artifacts (e.g., ocular, bad connection at a single channel) was performed semiautomatically through visual inspection. On average, 12.9 ICs (*SD* = 2.3) were rejected as artifacts.

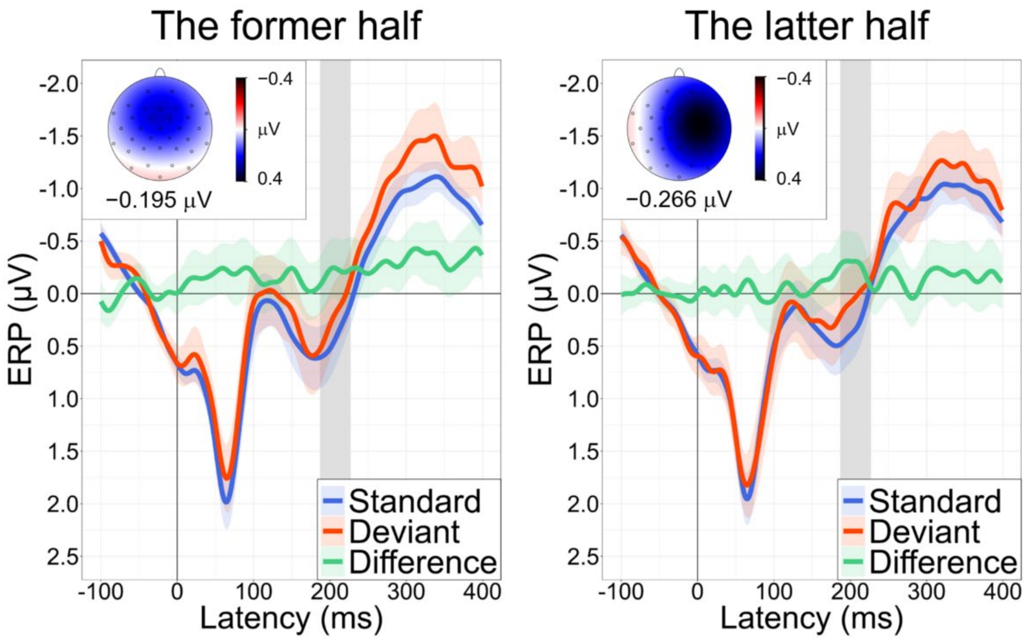
MMNs in the former and latter halves of the experiment

To examine whether learning affected the MMN amplitude or not, the MMN amplitudes were calculated separately from the first five (former half) and second five (latter half) blocks to examine the learning effect. Then, a mixed three-way analysis of variance (ANOVA) with condition (standard vs. deviance), block (former half vs. latter half), and group (Group I vs. Group II) was conducted on the ERP amplitude of the MMN interval. This analysis was also conducted using a Bayesian mixed three-way ANOVA to assess the absence (effect size $\delta = 0$, null hypothesis) or presence (effect size

$\delta \neq 0$, alternative hypothesis) of the effects. Furthermore, to examine the presence of the MMN, a one-sample *t*-test (one-sided) and its Bayesian analysis were conducted on the MMN amplitudes of the former and latter halves. Supplementary Figure S1 shows the grand average waveforms and scalp topographies of the ERPs elicited by the final chords of the former and latter halves. The mixed three-way ANOVA conducted on the MMN amplitudes revealed the significance of condition, $F(1, 34) = 8.057$, $p = .008$, $\eta_p^2 = .192$, $BF_{10} = 2.469$, suggesting that the MMN was elicited by the deviant chord transition irrespective of the combination of the chord. None of the other effects and interactions were significant, $F(1, 34) < 2.555$, $p > .119$, $\eta_p^2 < .070$, $BF_{10} < 0.578$. However, when the former half ($M = -0.195$ μV, $SD = 0.791$) and latter half ($M = -0.266$ μV, $SD = 0.701$) were analyzed separately, MMN amplitude was significantly negative in the latter half, $t(35) = -2.278$, $p = .014$, Cohen's $d = -0.380$, $BF_{-0} = 3.412$, but not in the former half, $t(35) = -1.480$, $p = .074$, Cohen's $d = -0.247$, $BF_{-0} = 0.892$. This finding can be seen as evidence of the learning effect, although the reliability of MMN measurements was lower than that of the original analysis using all 10 blocks due to a smaller number of averages.

Supplementary Figure S1.

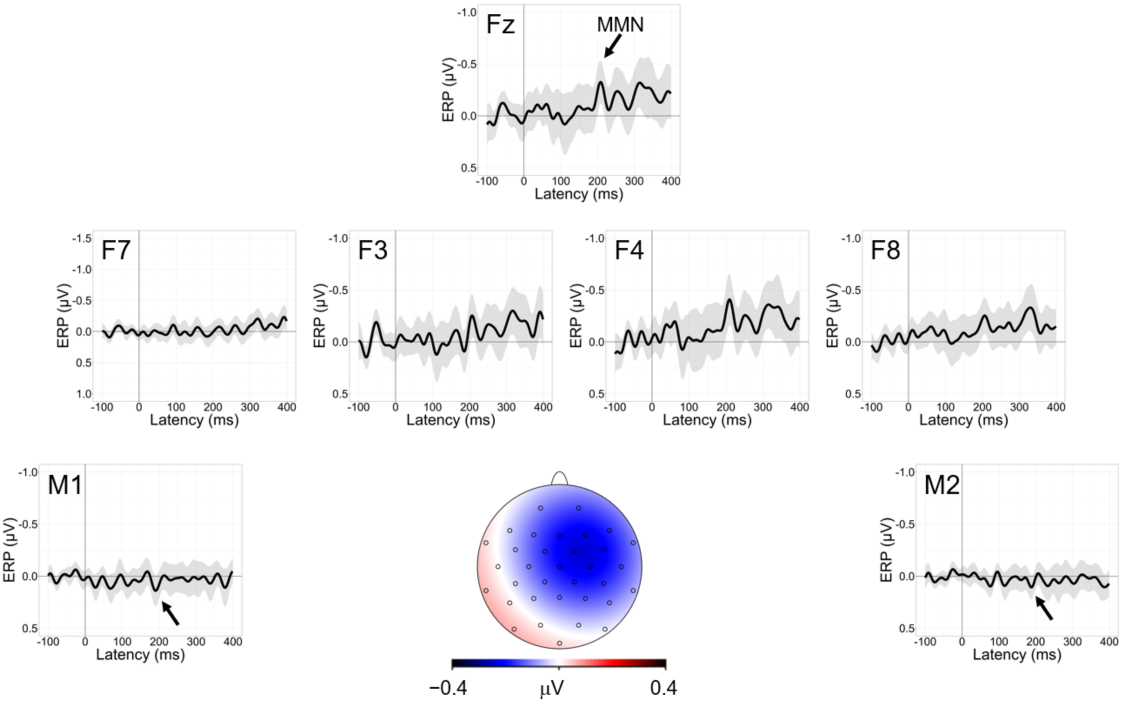
Grand average ERP waveforms and topography of the former and latter halves



Note. Grand average waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) with 95% confidence intervals and topographic maps (186–226 ms) of the original ERPs elicited by chords with high (standard) or low (deviant) transitional probability and deviant-related difference waveforms (difference) are shown.

Supplementary Figure S2.

Grand average difference waveforms of 5 frontal electrodes and the left and right mastoids



Note. ERP data were re-referenced to the nose. Difference waveforms were calculated by subtracting the ERP of chords with a high transitional probability from that of chords with a low transitional probability.