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	

1   **Statistical Learning of Chord-Transition Regularities in a Novel Equitempered**
2   **Scale: An MMN Study**
3

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

5   **REGULARITIES**

6

7   Kai Ishida[1], Hiroshi Nittono[1]

8   [1] Graduate School of Human Sciences, Osaka University, Osaka, Japan

9

10   **Corresponding author:** Kai Ishida, Graduate School of Human Sciences, Osaka

11   University, 1-2 Yamadaoka, Suita, Osaka 565-0871, JAPAN; ishida@hus.osaka-u.ac.jp;

12   ORCID: 0000-0001-6485-0950

13   Hiroshi Nittono, Graduate School of Human Sciences, Osaka University, 1-2

14   Yamadaoka, Suita, Osaka 565-0871, JAPAN; nittono@hus.osaka-u.ac.jp; ORCID:

15   0000-0001-5671-609X

16

1   The sound materials and dataset of the present paper are available at

2   https://osf.io/p725n/.

3

4   **CRediT author statement**

5   **Kai Ishida**: Conceptualization, Methodology, Investigation, Data curation, Formal

6   analysis, Visualization, Project administration, Funding acquisition, Writing–Original

7   Draft

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

9

10

11
12

1 **Highlights**

2 ● Chord sequences were created in a novel 18 equal temperament scale.

3 ● Chords were presented with high or low transitional probabilities.

4 ● Event-related potentials were recorded while participants listened to the chords.

5 ● The chords with low transitional probability elicited mismatch negativity.

6 ● Participants could not recognize the learned regularities beyond the chance level.

7

8 **Keywords**

9 statistical learning, mismatch negativity, event-related potential, musical chord

10 sequence, Markov process

11

1   **Abstract**

2   In music and language domains, it has been suggested that patterned transitions of

3   sounds can be acquired implicitly through statistical learning. Previous studies have

4   investigated the statistical learning of auditory regularities by recording early neural

5   responses to a sequence of tones presented at high or low transition probabilities.

6   However, it remains unclear whether the statistical learning of musical chord transitions

7   is reflected in endogenous, regularity-dependent components of the event-related

8   potential (ERP). The present study aimed to record the mismatch negativity (MMN)

9   elicited by chord transitions that deviated from newly learned transitional regularities.

10   Chords were generated in a novel 18 equal temperament pitch class scale to avoid

11   interference from the existing tonal representations of the 12 equal temperament pitch

12   class system. Thirty-six adults without professional musical training listened to a

13   sequence of randomly inverted chords in which certain chords were presented with high

14   (standard) or low (deviant) transition probabilities. An irrelevant timbre change

15   detection task was assigned to make them attend to the sequence during the ERP

16   recording. After that, a familiarity test was administered in which the participants were

17   asked to choose the more familiar chord sequence out of two successive sequences. The

18   results showed that deviant transitions elicited the MMN, although the participants

19   could not recognize the standard transition beyond the level of chance. These findings

20   suggest that humans can statistically learn new transitional regularities of chords in a

21   novel musical scale, even though they did not recognize them explicitly. This study

22   provides further evidence that music-syntactic regularities can be acquired implicitly

23   through statistical learning.

24

Introduction

1

2   In Western tonal music, each chord has a specific function (e.g., tonic, dominant

3 and subdominant), and the arrangement of chords follows the rule of harmony. Several

4 studies have shown that some aspects of tonal regularities in music can be acquired

5 through long-term learning and plasticity [1–2]. For example, an event-related potential

6 (ERP) study demonstrated that early right anterior negativity (ERAN), which reflects

7 the violation of harmonic expectancy [3–4], shows a greater amplitude for rarer

8 harmonic progressions based on the Western music corpus [5]. This result supports the

9 notion of probabilistic or statistical learning of transitional regularities through everyday

10 music listening.

11   Statistical learning is a process that realizes the grouping and segmentation of

12 events in various sensory modalities based on probabilistic regularity and has been

13 examined by behavioral and neural responses [6]. In the auditory domain, as evidence

14 that transitional patterns of tones can be learned statistically, magnetoencephalography

15 (MEG) studies have reported the attenuation of the magnetic counterparts of exogenous

16 ERPs such as P1 and N1 for tones with higher transitional probability than for tones

17 with lower transitional probability [7–8]. Moreover, recent studies have demonstrated

18 that endogenous mismatch negativity (MMN) is elicited by tones with low transitional

19 probability [9–12]. Because the MMN is the first endogenous component that reflects

20 memory-based prediction based on the regularity extracted from the preceding context,

21 the elicitation of the MMN can be a more direct indicator of statistical regularity

22 learning than the attenuation of exogenous components, which are more dependent on

23 sensory inputs.

24   Statistical learning has also been adopted for the acquisition of the regularity

1 representation of note transition patterns within a melody and of chord-transition

2 patterns [15–18]. Loui et al. [18] reported that infrequent chords in a novel musical

3 scale (i.e., Bohlen-Scale) elicited early anterior negativity with a latency of 150–210 ms.

4 They presented a chord progression in three different keys to enhance the

5 generalizability of the progression pattern. However, it is still possible that the

6 participants learned the patterns of pitch contours rather than the patterns of chord

7 progressions.

8   In a musical context, a chord is defined as a simultaneously sounded harmonic set

9 of tones with specific pitch classes rather than specific pitch heights. Even if two chords

10 consist of tones with different pitch heights, they are categorized as the same chord if

11 they consist of tones with the same pitch classes. For instance, the C major chord (C3–

12 E3–G3) is still categorized as the C major chord, even if component C3 is raised by one

13 octave (E3–G3–C4) and the chord is inverted. Considering this property of chords,

14 Daikoku et al. [15] conducted an MEG study in which a particular chord was repeated

15 three times in three different inversions and then transitioned to another chord with high

16 or low probabilities. The effect of statistical learning was observed as the attenuation of

17 an exogenous P1m for chords with high transitional probability compared to chords

18 with low transitional probability around 70 ms after chord onset. However, Daikoku et

19 al. did not examine the effect in a later latency range because they did not aim to record

20 later components such as the MMN.

21   The present study aimed to investigate whether the statistical learning of the

22 transitional regularity of chords is reflected in the preattentive MMN [13], the elicitation

23 of which would provide additional evidence for statistical learning in addition to the

24 attenuation of exogenous ERPs [15]. To avoid interference from existing tonal

1 representations, an 18-equal temperament scale was used, and six types of triad chords

2 were created. Triad chords were presented in three inversions to ensure that the target of

3 learning was a harmonic chord (i.e., a set of pitch classes) rather than a set of tones with

4 specific pitch heights [15]. To record the MMN, the experimental paradigm of Koelsch

5 et al. [9] and Tsogli et al. [12] was adopted. In their studies, statistical learning of

6 regularities, as indicated by the MMN, was implicit, because the participants were not

7 aware of the regularities. In the present study, the transitional regularities of chords were

8 manipulated. Various chord triplets were presented repeatedly without a pause. In each

9 chord triplet, the first two chords formed a "root," and the last chord was an "ending."

10 Each root transitioned to one of two types of endings at high ($p = .90$) or low ($p = .10$)

11 probabilities so that the same ending became either a standard or deviant transition,

12 depending on the roots. Because each chord was presented with equal probability and

13 the same chords would become either the standard or deviant, the MMN observed in the

14 present study could reflect deviant detection based on the regularity acquired through

15 statistical learning, rather than the processing of the occurrence frequency of chords or

16 the change in acoustic features. After the ERP experiment, the implicit or explicit nature

17 of learning was examined using a familiarity test to determine whether participants

18 could recognize the transitional regularities above the level of chance.

19 **Materials and Methods**

20 *Participants*

21 The sample size ($N = 34$) was determined using G*Power [19] to detect a medium

22 effect size ($dz = 0.5$) with a power of .80. This medium effect size was selected

23 according to the effect sizes reported in previous studies, which were often larger than

24 0.5 ($dz = 0.76$ [9] or 1.33 [12] for MMN; $dz = 0.69$ [18] for the early anterior

1   negativity). Expecting some dropouts, we recruited 36 adults without professional

2   musical training (23 women and 13 men, 18–38 years old, $M = 22.6$ years). However,

3   all participants' data could be used for hypothesis testing. All participants were right-

4   handed [20], and none had hearing impairments or a history of neurological disease.

5   The participants had various types of musical experience, with a mean of 5.8 years of

6   extracurricular musical lessons (range 0–17 years). The protocol was approved by the

7   Behavioral Research Ethics Committee of the Osaka University School of Human

8   Sciences, Japan (HB022-107), and written informed consent was obtained from all

9   participants. Participants received a cash voucher of 3,500 Japanese yen as an

10   honorarium.

11   Figure 1

12   *Materials*

13   Figure 1 shows all chords used in the present study. Six types of triad chords

14   consisting of notes from the 18 equal-temperament scale were created in three different

15   inversions while controlling interference by the Western music corpus [15]. The

16   rationale for using this scale was described in Supplementary Material. Each note was a

17   sine tone to avoid a timbre-specific effect. In the chord sequence, the duration of a

18   chord was 450 ms, which included a rise and fall of 10 and 200 ms. Each chord was

19   presented with an interstimulus interval of 50 ms (thus, the inter-onset interval was 500

20   ms). All chords were sampled at 44,100 Hz, and the amplitude was normalized.

21   All methods were similar to those in Tsogli et al. [12]. Three chords were connected

22   to form a triplet. Four types of triplets were created by combinations of two types of

23   roots (AC or BD) and two types of endings (E or F). Each root transitioned to either E

24   or F chords with high or low frequency. The left panel of Figure 2 shows how the

1   chords transitioned. To control for the combination effect of chords, whether the chords

2   were high or low transitions was counterbalanced between participants. In Group I, the

3   AC transitioned to the E chord with high probability ($p = .90$) and transitioned to the F

4   chord with low probability ($p = .10$), while the BD transitioned to the F chord with high

5   probability and transitioned to the E chord with low probability. In Group II, the AC

6   transitioned to the F chord with high probability and to the E chord with low probability,

7   while the BD transitioned to the E chord with high probability and to the F chord with

8   low probability. The transitional probability between the triplets was equal ($p = .50$).

9   Thus, all chords had an equal probability of occurrence. Note that three types of

10   inversions of a chord were randomly presented at each chord position.

11   *Procedure*

12   In the EEG recording, all four triplets were presented in random order.[1] During the

13   presentation, the participants were asked to press a button as quickly and accurately as

14   possible when the chords changed to a piano in a timbre change detection task. Using

15   this kind of detection task is a common method in statistical learning to make

16   participants pay attention to a stimulus sequence without focusing on its regularities [7–

17   9, 12, 15]. Timbre changes occurred at random positions within a triplet 3–5 times in

18   each block for a total of 40 changes in the entire experiment. The timbre change

19   occurred only in triplets including chords with high transitional probability to avoid

20   reducing the number of trials with low transitional probability. Ten blocks were

21   performed with short breaks, and each block lasted approximately six minutes. Within a

22   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.

1   with the constraint that triplets with low transitional probabilities were not repeated in

2   succession. In Group I, ACE, ACF, BDF, and BDE were presented 900, 100, 900, and

3   100 times, respectively. In Group II, ACF, ACE, BDE, and BDF were presented 100,

4   900, 100, and 900 times, respectively.

5   The EEG recording was followed by a two-alternative forced-choice familiarity test

6   that took four minutes. In the familiarity test, four types of possible pairs of the

7   unlearned triplets (ACE vs. BDE, BDE vs. ACE, ACF vs. BDF, and BDF vs. ACF) were

8   presented six times (i.e., 24 trials in total). The order of presentation of the roots was

9   counterbalanced across participants. The pause between two triplets of a pair was 500

10   ms. The participants' task was to choose which triplets sounded more familiar by

11   pressing a key that corresponded to either the first or second triplet. The choice of the

12   triplet that contained chords with high transitional probability was regarded as the

13   correct response. After choosing the triplet, participants described their confidence in

14   their choice using a scale from 1 = *very unsure* to 5 = *very sure* at their own pace. The

15   regularity of chord transitions was explained at the end of the experiment.

16   *EEG recording and data reduction*

17   EEG data were recorded using a QuickAmp (Brain Products) with Ag/AgCl

18   electrodes. Thirty-four scalp electrodes were applied according to the 10–20 system

19   (Fp1/2, F3/4, F7/8, Fz, FC1/2, FC5/6, FT9/10, C3/4, T7/8, Cz, CP1/2, CP5/6, TP9/10,

20   P3/4, P7/8, Pz, O1/2, Oz, PO9/10). Additional electrodes were placed on the left and

21   right mastoids, the left and right outer canthi of the eyes, and above and below the right

22   eye. The data were referenced offline to the algebraic means of the left and right

23   mastoid electrodes. The sampling rate was 1,000 Hz. The online filter was DC–200 Hz.

24   Electrode impedances were kept below 10 kΩ.

1   EEG data were analyzed using Brain Vision Analyzer (Brain Products, Germany).

2   First, a digital filter of 0.5 Hz (6 dB/oct) high-pass filter and 30 Hz (48 dB/oct) low-pass

3   filter and a notch filter of 60 Hz were applied to the data [9, 12]. After correcting ocular

4   and other artifacts (see Supplementary Material for details), a 500 ms period (100 ms

5   before and 400 ms after the ending note) was averaged after removing trials in which

6   voltages exceeded ±80 µV in any channel. Two consecutive trials after the timbre

7   change were removed from the analysis. Baseline correction was applied by subtracting

8   the mean amplitude of the prestimulus 100 ms from each point of the waveform.

9   Statistical evaluation was conducted at the frontal electrode cluster (F7, F3, Fz, F4, and

10 F8) based on the previous study [21]. The peak of MMN was detected in the interval of

11 150–280 ms of the grand mean difference waveform (using all 10 blocks), calculated by

12 subtracting the ERP waveform of chords with high transitional probability from that of

13 chords with low transitional probability. The interval ± 20 ms from the peak was defined

14 as the MMN interval. On average, 1644 (1496–1670) and 197 (177–200) epochs were

15 used to calculate the standard and deviant ERP waveforms, respectively.

16 *Statistical analysis*

17   Statistical analyses were carried out using JASP 0.17.2 [22]. A mixed two-way

18 analysis of variance (ANOVA) with condition (standard vs. deviance), and group

19 (Group I vs. Group II) was conducted on the ERP amplitude of the MMN interval. This

20 analysis was also conducted using a Bayesian mixed two-way ANOVA to assess the

21 absence (effect size $\delta = 0$, null hypothesis) or presence (effect size $\delta \neq 0$, alternative

22 hypothesis) of the effects. The correct percent of the familiarity test was aggregated

23 across the groups, and compared to the chance level ($p = .50$) using a one-sample *t*-test

24 (one-sided) because the correct percentages of both groups were not significantly

1   different, $t(31.5) = 0.114$, $p = .910$, Cohen's $d = .038$, $BF_{01} = 3.093$. The same analysis

2   was conducted using a Bayesian one-sample t-test to assess the absence (effect size $\delta = $

3   0, null hypothesis) or presence of the difference (effect size $\delta < 0$, alternative

4   hypothesis). Finally, the confidence rating between correct and incorrect responses was

5   compared using a paired t-test (two-sided) and a Bayesian paired t-test. For frequentist

6   hypothesis testing, the significance levels were set to $\alpha = .05$. For Bayesian hypothesis

7   testing, the Cauchy distribution with a scale parameter r of 0.707 was used as the prior

8   distribution for δ in the t-test. For the Bayesian two-way repeated-measures ANOVA,

9   multivariate Cauchy distribution (fixed effect: scale parameter $r = 0.5$; random effect:

10   scale parameter $r = 1$; covariates: scale parameter $r = .354$) was used as the prior

11   distribution. As an exploratory analysis, MMN amplitudes in the former and latter

12   halves of the experiment were compared to examine the effect of learning (see

13   Supplementary Material).

14

15   **Results**

16   The averaged mean reaction time of the timbre change detection task was 308 ms

17   ($SD = 46$ ms), and the averaged hit rate was 98.9% ($SD = 3.0\%$), suggesting that the

18   participants focused on the task and attended to the chord sequence.

19   Figure 2

20   The right panel of Figure 2 shows the grand average waveforms and scalp

21   topographies of the ERPs. Chords with low transitional probability elicited the MMN

22   ($M = -0.223$ μV, $SD = .487$) over the frontal area, with a peak latency of 206 ms.

23   Therefore, a period of 186–226 ms was used for scoring MMN amplitudes. Similar to

24   the typical MMN, a slight polarity inversion was observed when the data were

1   referenced to the nose (see Supplementary Figure S2). The mixed two-way ANOVA

2   conducted on the MMN amplitudes revealed the significance of condition, $F(1, 34) =$

3   7.331, $p = .011$, $\eta_p^2 = .177$, $BF_{10} = 4.532$, suggesting that the MMN was elicited by the

4   deviant chord transition irrespective of the combination of the chord. None of the other

5   effects and interactions were significant, $F(1, 34) < 0.368$, $p > .548$, $\eta_p^2 < .011$, $BF_{10} <$

6   0.561. Although the MMN was observed, the one-sample t-test showed that the

7   percentage of correct responses ($M = 52.5\%$, $SD = 12.0$) did not significantly exceed the

8   chance level, $t(35) = 1.277$, $p = .105$, Cohen's $d = 0.213$, $BF_{+0} = 0.671$. The difference

9   between confidence ratings when the response was correct ($M = 3.3$, $SD = 0.6$) or

10   incorrect ($M = 3.2$, $SD = 0.5$) was not significant, $t(35) = 0.741$, $p = .463$, Cohen's $d =$

11   0.124, $BF_{10} = 0.231$.

<div align="center">**Discussion**</div>

13   The present study examined whether the MMN response is elicited by deviations

14   from the statistically learned transitional regularity of chords, defined as a harmonic set

15   of pitch classes in a novel musical scale. The results of the ERP showed that a chord

16   elicited MMN when it was presented with a low transitional probability, even if the

17   chord was presented equiprobably in the whole experiment. The results of the

18   familiarity test, however, showed that the participants could not recognize the standard

19   transition beyond the level of chance, and there was no difference in confidence ratings

20   between correct and incorrect responses, suggesting that the participants chose the

21   triplets without clear response criteria. These behavioral results indicate that the

22   acquired representation was implicit.

23   This study provides further evidence that the transitional regularities of chords are

24   statistically learned by demonstrating that the MMN, which is a memory-based

1   endogenous component [14], is elicited by chords with low transitional probability.

2   Consistent with the MMN around 180–260 ms after the onset of the tone of Tsogli et al.

3   [12], the latency of the MMN response (186–226 ms) in the present study was later than

4   that of the traditional MMN response, such as a physical MMN elicited by the

5   infrequent change in the acoustic feature of tone (e.g., 100–200 ms; as reviewed by 13–

6   14). Tsogli et al. suggested that the generation of the statistical MMN requires more top-

7   down processing to encode the deviance than the physical MMN because a longer time

8   is needed to learn the contextual regularity. This coincides with previous findings that

9   the latency of the MMN is longer as the complexity of stimuli increases [23]. Moreover,

10   the exploratory analysis showed that a significant MMN occurred only in the latter half

11   of the experiment, which supports the learning effect (see Supplementary Material).

12   The MMN response of the present study and the early anterior negativity reported

13   by Loui et al. [18] may be the same kind of component that reflects auditory deviant

14   detection based on the statistical learning of chord progressions or transition

15   regularities. By manipulating chord inversions, the present study extends the findings of

16   Loui et al. to the more abstract harmonic regularity, where chords are defined as a

17   harmonic set of pitch classes in a novel musical scale. Furthermore, the generation

18   process of the MMN in the current study may be similar to that of the ERAN. The

19   ERAN reflects the violation of harmonic expectancy based on the schema of musical

20   syntax acquired as a long-term format [2]. MMN is thought to reflect the innate ability

21   to extract regularities in the relationship between sounds and immediately establish

22   representations, as MMN-like discriminative responses are elicited by changes in

23   acoustic features from the infant stage [24]. Through a review of ERAN studies,

24   Koelsch [2] noted that the formation and organization processes of auditory objects,

1   which are required to generate the MMN, are indispensable for conducting music-

2   syntactic processing. Furthermore, some studies have proposed that the ERAN and

3   MMN reflect similar irregularity detection processes, despite differences in regularity

4   representations [4, 21]. In the present study, the learning processes of chord transitions

5   may be similar to the process required to acquire the pattern representations of harmony.

6   Taken together, the regularity representations underlying music-syntactic processing are

7   possibly acquired by statistical or probabilistic learning [2, 5].

8   In the present study, the participants could not recognize the regular transition

9   pattern beyond the chance level. Previous studies have also reported that the MMN can

10   be elicited when the performance on the follow-up behavioral test is below the chance

11   level [9, 11–12]. ERPs have been considered a more sensitive measure of statistical

12   learning than behavioral measures [25]. It is also possible that the discrepancy between

13   neural and behavioral results may reflect insufficient learning. The MMN amplitude of

14   the present study was smaller than that of the early anterior negativity of Loui et al.

15   [18], in which explicit recognition beyond the chance level was achieved. The regularity

16   in the present musical stimuli might be difficult to learn sufficiently in one hour of

17   listening. Moreover, musical proficiency may facilitate statistical learning [8]. Because

18   the present study did not control for participants' musical experience and absolute pitch

19   ability, future research is needed to examine the relationship between ERP and

20   behavioral measures in statistical learning of musical stimuli.

21   In summary, the present results demonstrated that chords with low transitional

22   probability elicited the MMN. This is due to the statistical learning of transitional

23   regularities of chords (defined as a harmonic set of pitch classes) in a novel musical

24   scale. The participants could not recognize the standard transition chords beyond the

1   level of chance. Future neuroscientific research should examine whether explicit

2   knowledge of regularity can be acquired when the regularity is learned intentionally. In

3   conclusion, the present study suggests that the representation of music-syntactic

4   regularities can be acquired through statistical learning.

1   **References**

2   [1] H. Honing, C. ten Cate, I. Peretz, & S. E. Trehub, Without it no music: Cognition,

3   biology and evolution of musicality. Philos. Trans. R. Soc. 370 (2015) 20140088.

4   https://doi.org/10.1098/rstb.2014.0088

5   [2] S. Koelsch, Music-syntactic processing and auditory memory: Similarities and

6   differences between ERAN and MMN. Psychophysiology 46 (2009) 179–190.

7   https://doi.org/10.1111/j.1469-8986.2008.00752.x

8   [3] S. Koelsch, T. Gunter, A. D. Friederici, & E. Schröger, Brain indices of music

9   processing: "Nonmusicians" are musical. J. Cogn. Neurosci. 12 (2000) 520–541.

10   https://doi.org/10.1162/089892900562183

11   [4] S. Koelsch, T. C. Gunter, E. Schröger, M. Tervaniemi, D. Sammler, & A. D.

12   Friederici, Differentiating ERAN and MMN: An ERP study. NeuroReport 12

13   (2001) 1385–1389. https://doi.org/10.1097/00001756-200105250-00019

14   [5] S. G. Kim, J. S. Kim, & C. K. Chung, The effect of conditional probability of chord

15   progression on brain response: An MEG study. PLOS ONE 6 (2011) e17337.

16   https://doi.org/10.1371/journal.pone.0017337

17   [6] C. M. Conway, How does the brain learn environmental structure? Ten core

18   principles for understanding the neurocognitive mechanisms of statistical learning.

19   Neurosci Biobehav Rev. 112 (2020) 279–299.

20   [7] T. Daikoku, Y. Yatomi, & M. Yumoto, Implicit and explicit statistical learning of

21   tone sequences across spectral shifts. Neuropsychologia 63 (2014) 194–204.

22   https://doi.org/10.1016/j.neuropsychologia.2014.08.028

23   [8] T. Daikoku, & M. Yumoto, Musical expertise facilitates statistical learning of

24   rhythm and the perceptive uncertainty: A cross-cultural study. Neuropsychologia,

1   146 (2020) 107553. https://doi.org/10.1016/j.neuropsychologia.2020.107553

2   [9] S. Koelsch, T. Busch, S. Jentschke, & M. Rohrmeier, Under the hood of statistical

3   learning: A statistical MMN reflects the magnitude of transitional probabilities in

4   auditory sequences. Sci. Rep. 6 (2016) 19741. https://doi.org/10.1038/srep19741

5   [10] T. Moldwin, O. Schwartz, & E. S. Sussman, Statistical learning of melodic patterns

6   influences the brain's response to wrong notes. J. Cogn. Neurosci. 29 (2017) 2114–

7   2122. https://doi.org/10.1162/jocn

8   [11] E. Paraskevopoulos, A. Kuchenbuch, S. C. Herholz, & C. Pantev, Statistical

9   learning effects in musicians and non-musicians: An MEG study.

10   Neuropsychologia 50 (2012) 341–349.

11   https://doi.org/10.1016/j.neuropsychologia.2011.12.007

12   [12] V. Tsogli, S. Jentschke, T. Daikoku, & S. Koelsch, When the statistical MMN

13   meets the physical MMN. Cogn. Sci. 9 (2019) 5563.

14   https://doi.org/10.1038/s41598-019-42066-4

15   [13] R. Näätänen, T. Jacobsen, & I. Winkler, Memory-based or afferent processes in

16   mismatch negativity (MMN): A review of the evidence. Psychophysiology 42

17   (2005) 25–32. https://doi.org/10.1111/j.1469-8986.2005.00256.x

18   [14] E. S. Sussman, S. Chen, J. Sussman-Fort, & E. Dinces, The five myths of MMN:

19   Redefining how to Use MMN in basic and clinical research. Brain Topogr. 27

20   (2014) 553–564. https://doi.org/10.1007/s10548-013-0326-6

21   [15] T. Daikoku, Y. Yatomi, & M. Yumoto, Pitch-class distribution modulates the

22   statistical learning of atonal chord sequences. Brain Cogn. 108 (2016) 1–10.

23   https://doi.org/10.1016/j.bandc.2016.06.008

24   [16] E. M. M. Jonaitis, & J. R. Saffran, Learning harmony: The role of serial statistics.

1 Cogn. Process. 33 (2009) 951–968. https://doi.org/10.1111/j.1551-
2 6709.2009.01036.x

3 [17] P. Loui, Learning and liking of melody and harmony: Further studies in artificial

4 grammar learning. Top. Cogn. Sci. 4 (2012) 554–567.

5 https://doi.org/10.1111/j.1756-8765.2012.01208.x

6 [18] P. Loui, E. H. Wu, D. L. Wessel, & R. T. Knight, A generalized mechanism for

7 perception of pitch patterns. J. Neurosci. 29 (2009) 454–459.

8 https://doi.org/10.1523/JNEUROSCI.4503-08.2009

9 [19] F. Faul, E. Erdfelder, A. G. Lang, & A. Buchner, G* Power 3: A flexible statistical

10 power analysis program for the social, behavioral, and biomedical sciences.

11 Behavior Research Methods, 39 (2007) 175–191.

12 https://doi.org/10.3758/bf03193146

13 [20] R. C. Oldfield, The assessment and analysis of handedness: the Edinburgh

14 inventory. Neuropsychologia 9 (1971) 97-113. https://doi.org/10.1016/0028-

15 3932(71)90067-4

16 [21] K. Ishida, & H. Nittono, Relationship between early neural responses to syntactic

17 and acoustic irregularities in music. Eur. J. Neurosci. 56 (2022) 6201–6214.

18 https://doi.org/10.1111/ejn.15856

19 [22] JASP Team, JASP (version 0.17.2) [computer software] (2023). https://jasp-

20 stats.org/.

21 [23] T. Kujala, M. Tervaniemi, & E. Schröger, The mismatch negativity in cognitive

22 and clinical neuroscience: Theoretical and methodological considerations. Biol.

23 Psychol1. 74 (2007) 1–19. https://doi.org/10.1016/j.biopsycho.2006.06.001

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

1   for investigating auditory discrimination and sensory memory in infants and

2   children. Clin. Neurophysiol. 111 (2000) 4–16. https://doi.org/10.1016/S1388-

3   2457(99)00191-1

4   [25] T. Okano, T. Daikoku, Y. Ugawa, K. Kanai, & M. Yumoto, Perceptual uncertainty

5   modulates auditory statistical learning: A magnetoencephalography study. *Int J*

6   *Psychophysiol.* 168 (2021) 65–71. https://doi.org/10.1016/j.ijpsycho.2021.08.002

7

8

9

10

11

Figure 1.

Chords used in the present study and the pitch helix of 18 equal temperament
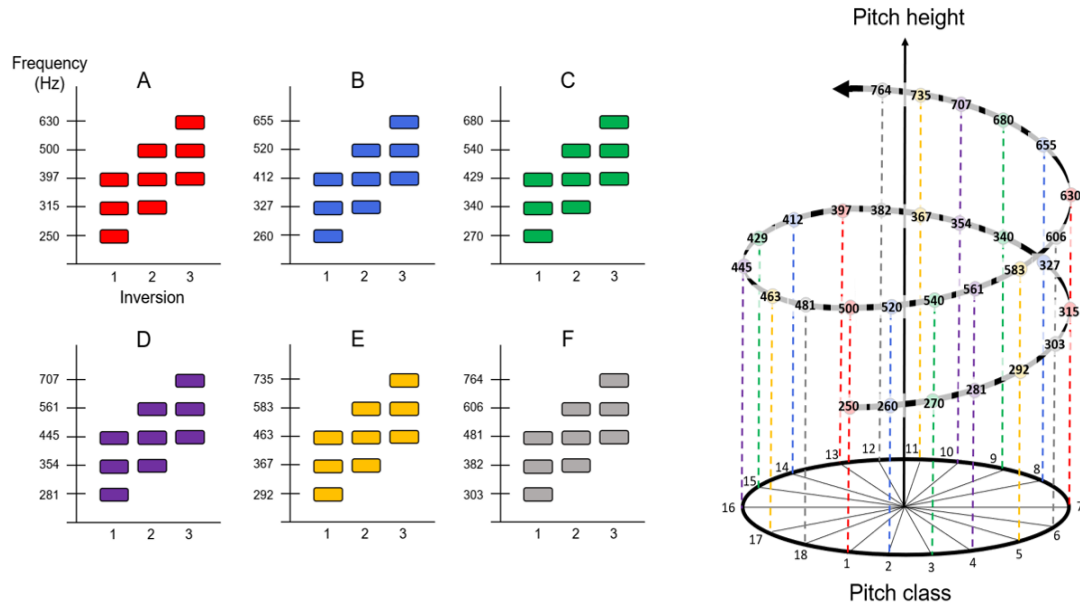
Note. The left panel shows the six types of triad chords and their three inversions. The vertical axis indicates the frequency (Hz), and the horizontal axis indicates the versions of chord inversion. For example, chord A could be [250, 315, 397 Hz], [315, 397, 500 Hz], or [397, 500, 630 Hz]. The right 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
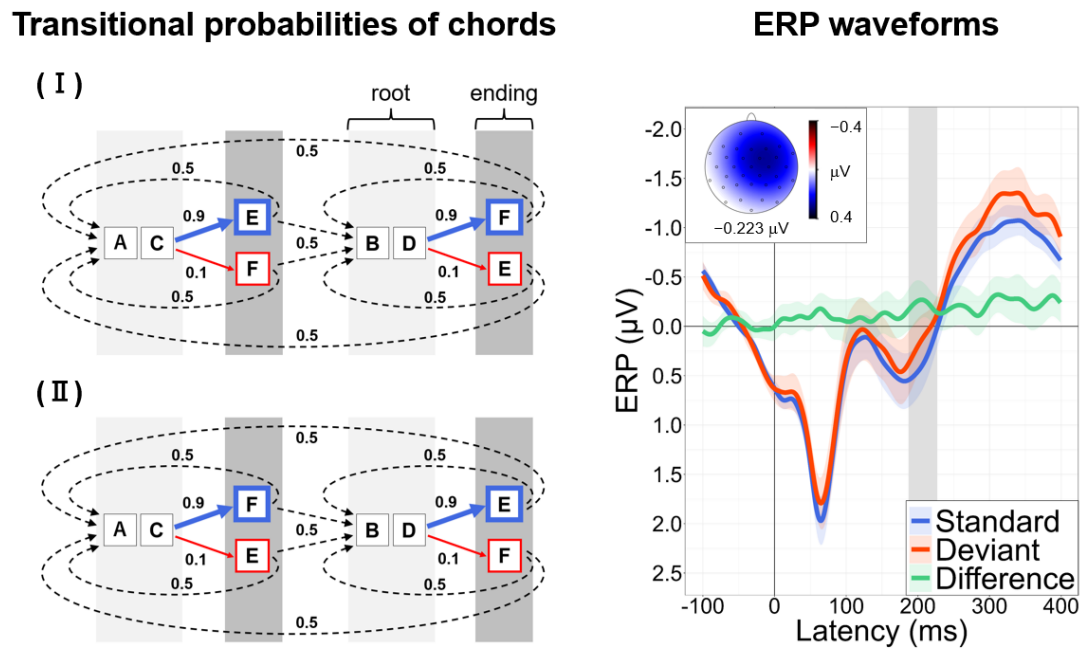
Note. In the left panel, Roman numerals in parentheses of the top and bottom figures indicate the group. In the right panel, grand average waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) with 95% confidence intervals and the topographic map (186–226 ms) of the deviant-related 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

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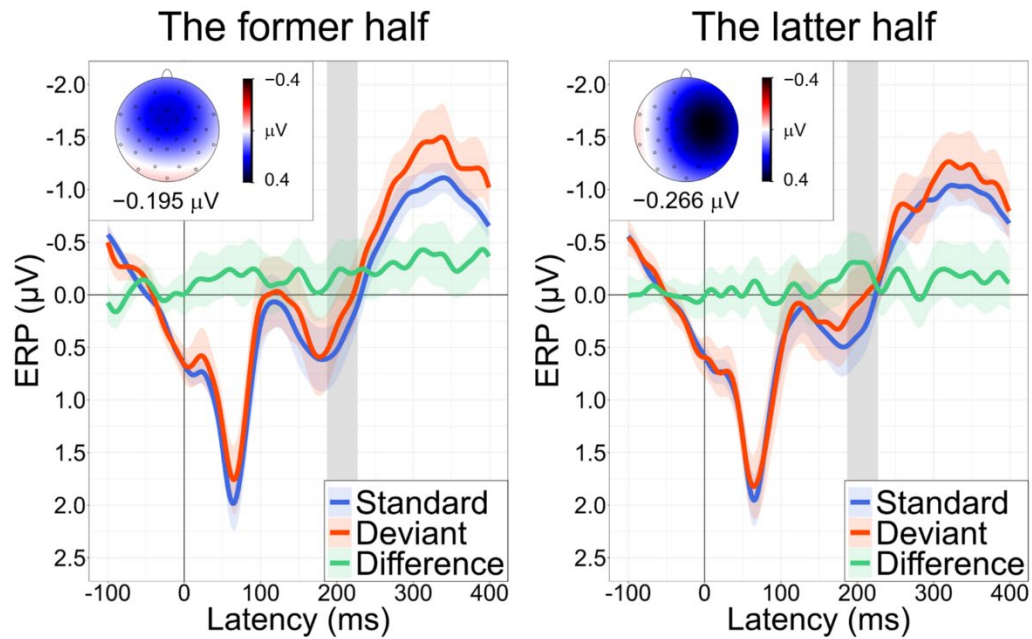
Note. Grand average waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) with 95% confidence intervals and topographic maps (186–226 ms) of the original ERPs elicited by chords with high (standard) or low (deviant) transitional probability and deviant-related difference waveforms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

Supplementary Figure S2.

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

mastoids
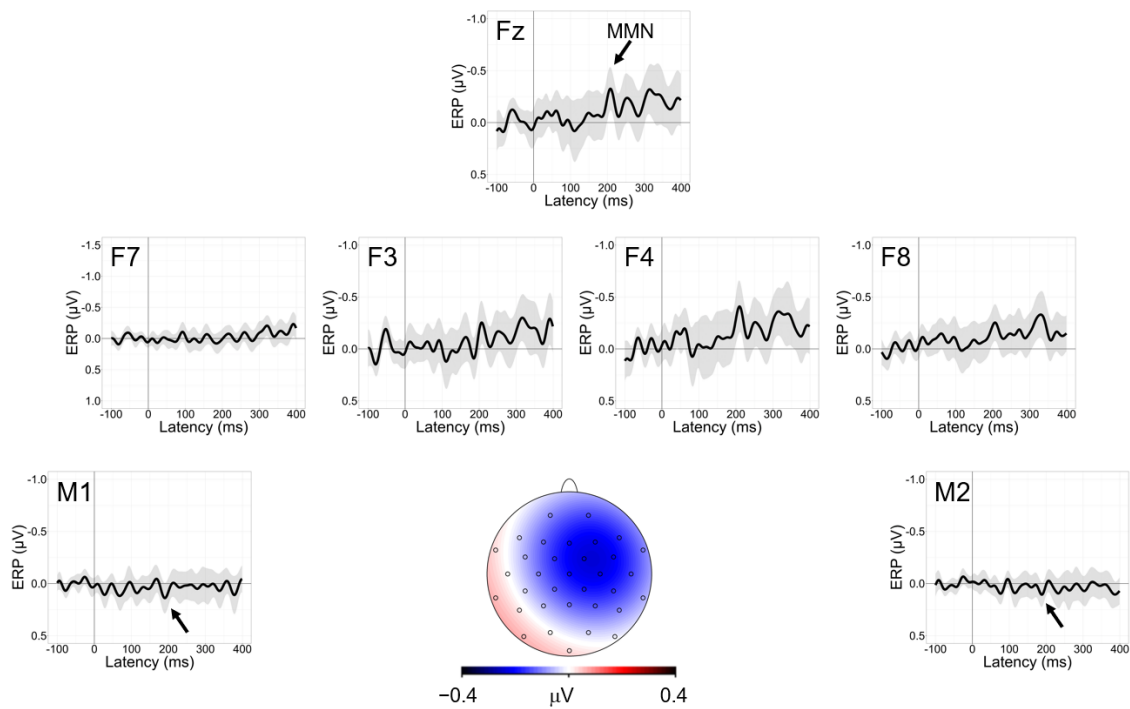
Note. ERP data were re-referenced to the nose. Difference waveforms were calculated

by subtracting the ERP of chords with a high transitional probability from that of chords

with a low transitional probability.