Neuropsychologia 204 (2024) 109012

Contents lists available at ScienceDirect

Neuropsychologia

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

Different plasticity patterns of schematic and dynamic expectations in musical pitch prediction

Kai Ishida[a,b,*], Hiroshi Nittono[a,**]

[a] Graduate School of Human Sciences, Osaka University, Osaka, Japan
[b] Japan Society for the Promotion of Science, Tokyo, Japan

1. Introduction

Predictive processes in music have been widely acknowledged (for a review, see Rohrmeier and Koelsch, 2012; Vuust et al., 2022). Electrophysiological studies have found differences in musical expectation dimensions by recording event-related potentials (ERPs) as a reflection of prediction errors (Koelsch et al., 2019). Mismatch negativity (MMN) is a deviance detection response to an irregularity in auditory patterns extracted from the current auditory context (Garrido et al., 2009; Näätänen et al., 2005; Sussman et al., 2014). It has been recorded not only for simple auditory stimuli but also in musical contexts, such as when the pitch of the melody is mistuned (Brattico et al., 2006), rhythmic patterns are unexpectedly changed (Vuust et al., 2009), and minor chords are presented occasionally in a sequence of major chords (Virtala et al., 2011). MMN elicitation depends on online regularity extraction, and the expectation reflected in the MMN can be dynamically updated (Garrido et al., 2009; Sussman & Winkler, 2001). For example, Lieder et al. (2013) proposed that the MMN reflects trial-by-trial statistical learning and showed that the MMN of each trial is influenced by the preceding tone sequence, and that a free-energy based Bayesian information processing model explained this history effect better than other plausible models (e.g., change detection hypothesis and adaptation hypothesis). As reflected in MMN elicitation, music prediction generates dynamic expectations that are immediately formed based on repeated patterns in a musical piece.

Studies of music-syntactic processing have dissociated early right anterior negativity (ERAN) from the MMN response (Ishida and Nittono, 2022; Koelsch et al., 2001; Leino et al., 2007; for a review, see Koelsch, 2009). ERAN is elicited by music-syntactic irregularity, such as irregularity in the rule of harmony (Koelsch et al., 2000) and out-of-key notes that do not belong to the key to the melody (Miranda and Ullman, 2007). Harmonically irregular chords presented with the same probability as regular chords can elicit ERAN (Ishida and Nittono, 2022; Koelsch et al., 2000, 2013), and ERAN can be observed even when the sensory dissonance of out-of-key notes is controlled (Koelsch et al., 2007; Koelsch and Sammler, 2008). As reflected in ERAN elicitation, music prediction generates schematic expectations that depend on musical knowledge acquired through long-term learning (Vuust et al., 2022).

In an ERP study, Ishida and Nittono (2024) examined the relationship between schematic and dynamic expectations. In their study, music-syntactic irregularity (i.e., violation of schematic expectations) and contour deviance (i.e., violation of dynamic expectations) occurred independently or simultaneously at the final position of melodies. The results showed that music-syntactic irregularity elicited the ERAN as a prediction error in schematic expectation, and contour deviance elicited the MMN as a prediction error in dynamic expectation. Furthermore, the ERP amplitude was multiplied when irregularity and deviance occurred simultaneously. Thus, Ishida and Nittono determined that schematic and dynamic expectations may operate interactively, with each expectation being updated separately. This interpretation is reasonable considering that ERAN and MMN are dissociable components (Ishida and Nittono, 2022; Koelsch et al., 2001; Leino et al., 2007).

Previous studies on predictive processing in music perception have categorized multiple types of expectations and examined the relationships between them (Guo and Koelsch, 2016; Ishida and Nittono, 2022, 2024; Tillmann and Bigand, 2010; for a review, see Vuust et al., 2022). However, little is known about how these expectations differ from each other in terms of their updating processes.

The purpose of the present study was to empirically compare the plasticity of schematic and dynamic expectations when each expectation simultaneously predicts the pitch class dimension by replicating and extending Ishida and Nittono's (2024) experimental paradigm. Music-syntactic irregularity and melodic-contour deviance (i.e., rare endings in the current auditory context) were presented independently or simultaneously at the final position of the melody, where the leading note induced an expectation for the tonic note. In addition, the

* Corresponding author. Graduate School of Human Sciences, Osaka University, 1-2 Yamadaoka, Suita, Osaka, 565, Japan.
** Corresponding author. Graduate School of Human Sciences, Osaka University, 1-2 Yamadaoka, Suita, Osaka, 565, Japan.
E-mail addresses: ishida@hus.osaka-u.ac.jp (K. Ishida), nittono@hus.osaka-u.ac.jp (H. Nittono).

https://doi.org/10.1016/j.neuropsychologia.2024.109012
Received 24 April 2024; Received in revised form 26 September 2024; Accepted 7 October 2024
Available online 9 October 2024

transitional probabilities of the final notes (i.e., the frequencies of the melodic contour patterns) were reversed in the first and second sessions of the experiment. Given that the underlying representation of schematic expectations is acquired through long-term exposure to music, the predictive model of schematic expectations should be more stable than that of dynamic expectations, suggesting that there is less plasticity for schematic expectations than for dynamic expectations. Therefore, we examined whether the ERAN and MMN amplitudes were differently modulated after the contextual change. Subjective fitness ratings were obtained to confirm whether the violations of expectations were recognized as deviations.

A number of hypotheses were formulated in the present study. First, it was predicted that the ERAN and MMN would occur in a similar latency range (H1). Second, the interactive effect of syntactic irregularity and contour deviance on these deviance-related potentials was expected in the first session, as Ishida and Nittono (2024) reported (H2). Third, if schematic expectations were more stable than dynamic expectations, syntactically irregular notes would elicit the ERAN across the sessions regardless of the transitional probabilities (H3a), whereas contour deviants in the first session would not elicit the MMN in the second session when the transitional probabilities of endings changed (H3b). Here, we expected two possible consequences of the probability change: MMN elicitation for novel rare stimuli (if expectation updating occurs rapidly) or MMN extinction (if expectation has inertia from the first session). Fourth, syntactically irregular notes would be rated less fit than syntactically regular notes in the first session (H4a), and contour deviant notes would be rated less fit than standard notes in the first session (H4b).

2. Methods

This study was preregistered before sampling. The preregistration details for the experiment can be found at https://osf.io/u6svg.

2.1. Participants

To ensure detection of a medium effect size ($d_z = 0.51$) for the interaction of music-syntactic irregularity and melodic-contour deviance, calculated by F (1, 36) = 9.76 (Ishida and Nittono, 2024) with power $1-\beta = 0.80$ and error rate $\alpha = .05$, a sample size of 33 was predetermined using G*power (Faul et al., 2007). A total of 40 participants were recruited, taking into account data exclusions. Participants were randomly assigned to two groups (Group I and Group II) with different probabilities of irregularities (see Stimulus section for details) to counterbalance the combination effects of irregularities. Ultimately, data from 36 participants (17 women and 19 men, 19–65 years old, $M = 23.9$ years, equal number of participants for each group) were used for hypothesis testing after three participants were excluded with excessive noise. Of these, 33 participants were right-handed, one was left-handed, and two were ambidextrous according to the FLANDERS handedness questionnaire (Okubo et al., 2014). None of the participants had hearing impairments or a history of neurological disease. They had a mean of 3.2

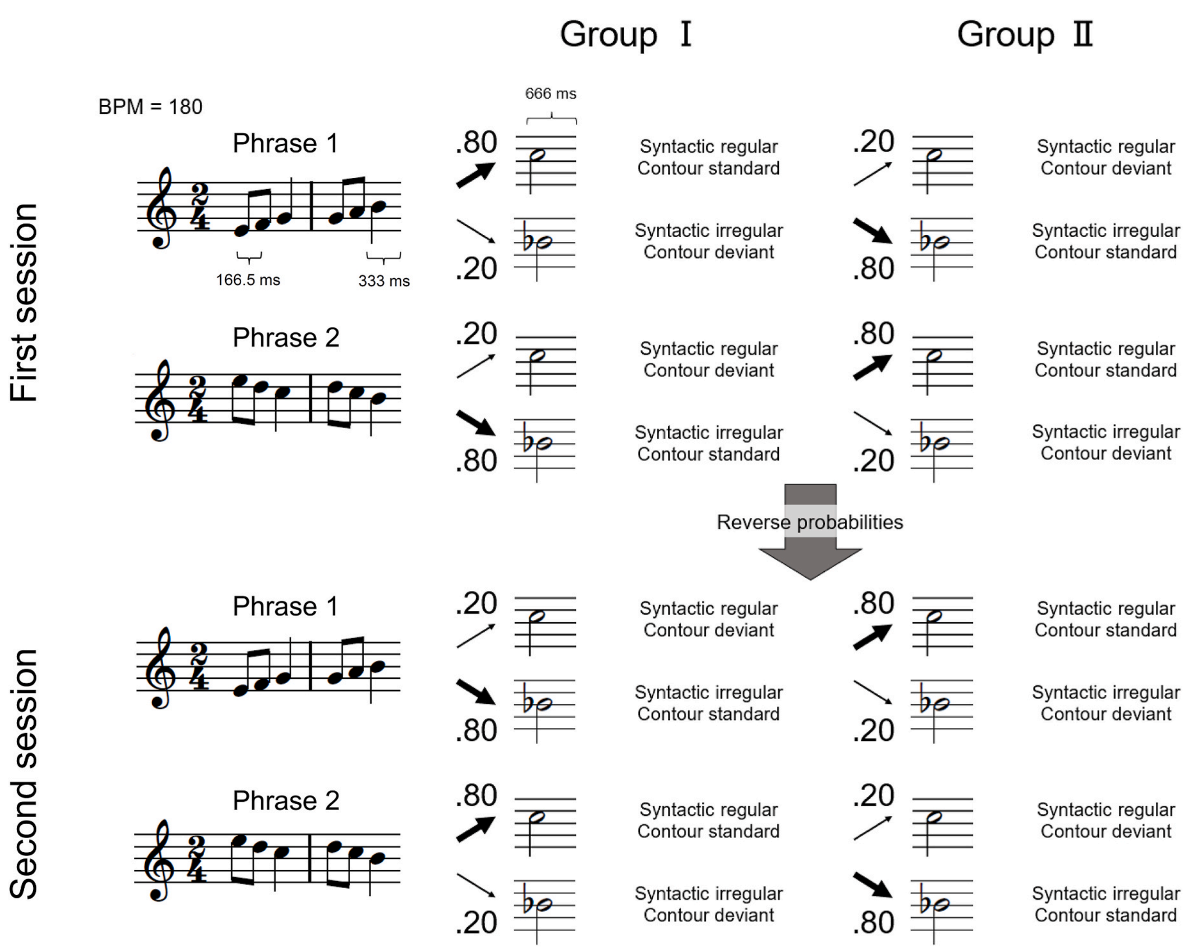

Fig. 1. Combinations of phrases and final notes in the first and second sessions.
Note. Four types of melodies consisted of two phrases and two final notes. Two final notes were syntactically regular or irregular to manipulate music-syntactic irregularity. The transitional probabilities of two phrases to two final notes were high ($p = .90$, standard) or low ($p = .10$, deviant) to manipulate contour regularity. The transitional probabilities were reversed in the second session. Combinations of syntactic and contour patterns were counterbalanced between groups.

(SD = 4.2) years of extracurricular musical lessons (range 0–15 years) and a mean of 0.6 (SD = 1.7) years of musical theory lessons including solfege (range 0–6 years). The protocol of this study was approved by the Behavioral Research Ethics Committee of the Osaka University School of Human Sciences, Japan (HB023-173R2), and written informed consent was obtained from all participants. Participants received a cash voucher of 2,500 Japanese yen as an honorarium.

2.2. Stimuli

Fig. 1 shows an example of the present stimuli. Two phrases were created using different combinations of the four measures (A, B, C, and D) used by Ishida and Nittono (2024). Each measure consisted of two eighth notes (166.5 ms each) and one quarter note (333 ms). While the previous study used AC and BD combinations, this study used AD and BC combinations for Phrases 1 and 2. They were then combined with a final note to form a melody. At the final position of a melody, a half note (666 ms) was music-syntactic regular (e.g., C in C major) or irregular (e.g., B♭ in C major) with high (p = .80, standard) or low (p = .20, deviant) probability. The melodies were played with a piano timbre.

Table 1 shows the probabilities of the melodies. Both syntactically regular and irregular notes were one semitone higher and lower than the preceding note, and the directions were different (i.e., up or down). Thus, four types of melodies (two phases × two final notes) were presented to each participant. The effect of melodic contour was controlled by setting different transitional probabilities for Group I and Group II. Moreover, the transitional probabilities were reversed between the first and second sessions. In the first session of Group I, Phrase 1 transitioned to the syntactically regular note with high probability and to the syntactically irregular note with low probability, whereas Phrase 2 transitioned to the syntactically irregular note with high probability and to the syntactically regular note with low probability. In the second session, Phrase 1 transitioned to the syntactically irregular note with high probability and to the syntactically regular note with low probability, whereas Phrase 2 transitioned to the syntactically regular note with high probability and to the syntactically irregular note with low probability. In Group II, the syntactically regular and irregular notes were presented with the opposite probability to Group I (see also Fig. 1 and Table 2).

Each melody was connected without any interstimulus interval. Because the transitional probabilities between melodies were equal (p = .50), the syntactically regular and irregular notes had the same probability of occurrence at the final position as in previous ERAN studies (Koelsch et al., 2007). In addition, each of the four melodies in each of the 12 major key versions (48 stimuli in total) was presented several times in randomized order to avoid the effect of sensory dissonance due to out-of-key notes in the syntactic irregular condition. However, because Phrase 1 contained the same standard notes repeatedly (e.g., two C notes in C major key) and this might cause sensory adaptation, we compared the ERAN responses between Phrase 1 and Phrase 2 (see Supplementary Materials, Fig. S1). The results showed that the ERAN responses were similar between phrases, suggesting that the ERAN response was elicited independently of sensory adaptation.

2.3. Procedure

During the EEG recording, the participants' task was to press two buttons simultaneously with both thumbs, as quickly and accurately as possible, when the melodic note changed from a piano timbre to a guitar timbre, as in Ishida and Nittono (2024). They also watched a silent movie to keep their eyes open and prevent them from falling asleep. The timbre changes occurred randomly 3–5 times in each melody sequence (i.e., one block). A total of 40 changes occurred in the experiment. To avoid reducing the number of trials with deviant conditions, the timbre changes occurred only in the final note with a high transitional probability. Ten blocks, each lasting approximately 6 min, were performed with short breaks. The transitional probabilities were reversed between the first five blocks (i.e., first session) and the second five blocks (i.e., second session). Within a block, Phrases 1 and 2 were presented 100 times each, and the syntactically regular and irregular notes were presented 100 times each, with the constraint that melodies with low transitional probability were not repeated consecutively. Table 2 shows the number of melodies presented in each session. The standard, syntactic irregular, contour deviant, and double-deviant conditions were presented with 400, 400, 100, and 100 trials, respectively, in the first and second sessions for both groups.

Table 1
Distribution of music-syntactic irregularity and transitional probabilities in the final notes of melodies.

	Musical syntax	
Melodic contour	Regular (0.50)	Irregular (0.50)
Standard (0.80)	0.40	0.40
Deviant (0.20)	0.10	0.10

Table 2
Number of trials for each type of final note occurring in each session.

Group	Session	Phrase	Final note	
			Syntactic regular	Syntactic irregular
I	First	1	400	100
		2	100	400
	Second	1	100	400
		2	400	100
II	First	1	100	400
		2	400	100
	Second	1	400	100
		2	100	400

The fitness rating task was conducted twice as a manipulation check: once after the first session and once after the second session. In this task, four types of melodies were presented once each in the 12 major key variations (48 stimuli in total), and the participants' task was to rate the fitness of the final note of the melody on a scale from 1 (*not fit at all)* to 7 (*very fit*). The same final note was not used in consecutive trials, and the presentation order of the phrases was counterbalanced. After the fitness rating task, debriefing about the regularity of the melodies and the presence of the syntactic irregularity and contour deviance was conducted.

2.4. EEG recording

EEG recordings were conducted using a QuickAmp (Brain Products, Germany) with Ag/AgCl electrodes. Thirty-four scalp electrodes were placed 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), and reference electrodes were placed on the left and right mastoids. To record the electrooculogram, additional electrodes were placed on the left and right outer canthi of the eyes and above and below the right eye. After recording, the data were referenced offline to the algebraic means of the left and right mastoid electrodes. The sampling rate was 1000 Hz. The online filter was DC–200 Hz. Electrode impedances were kept below 10 kΩ.

2.5. EEG data reduction

EEG data were analyzed using a Brain Vision Analyzer (Brain Products, Germany). First, a band-pass digital filter of 0.25 Hz (6 dB/oct)–25 Hz (48 dB/oct) was applied to the data (Koelsch et al., 2007). One bad electrode was corrected for each of the four participants by topographic interpolation (spherical interpolation). Ocular artifact correction was then applied based on independent component analysis using the Infomax algorithm. The ICs of artifacts (e.g., ocular, and bad connection on a single channel) were semiautomatically detected by visual inspection.

On average, 13.2 ICs (standard deviation = 2.4) were rejected as artifacts. The data were segmented into a 500 ms period (100 ms before and 400 ms after the final note), and trials with voltages exceeding ±80 μV in any channel were removed. The valid trials were then averaged. Two consecutive trials after the timbre change were removed from the averaging. Baseline correction was performed by subtracting the mean amplitude of the prestimulus 100 ms period from each point of the waveform.

Five frontal electrodes (F7, F3, Fz, F4, and F8) were clustered for statistical evaluation. First, to determine the latency windows for statistical analysis, the ERAN and MMN were calculated across two sessions using the whole ten blocks. The ERAN was calculated by subtracting the ERP waveforms elicited by music-syntactic regular notes from the ERP waveforms elicited by irregular notes. The MMN was calculated by subtracting the ERP waveforms elicited by frequent contour patterns from the ERP waveforms elicited by rare contour patterns. The ERAN (90–130 ms) and MMN (133–173 ms) time windows were defined as the 40 ms period centered around the most negative peak between 100 and 300 ms on each difference waveform.

Second, using these time windows, the ERAN and MMN amplitudes were calculated separately in the first and second sessions to examine whether the deviants in the first session elicited the ERAN and MMN in the second session when the transitional probabilities were reversed. The ERAN in each session was calculated by subtracting the ERP waveforms elicited by music-syntactic regular notes from the ERP waveforms elicited by irregular notes, irrespective of contour patterns. The MMN in the first session was calculated by subtracting the ERP waveforms elicited by frequent contour patterns from the ERP waveforms elicited by rare contour patterns, irrespective of their syntactic regularity. Conversely, the MMN in the second session was calculated by subtracting the ERP waveforms elicited by rare contour patterns (i.e., frequent contour patterns in the first session) from the ERP waveforms elicited by frequent contour patterns (i.e., rare contour patterns in the first session).

Finally, for the simultaneous analysis of the ERAN and MMN components, a single common time window was defined as the 40 ms period centered around the most negative peak between 100 and 300 ms on the average waveform of the ERAN and MMN grand mean waveforms. Based on this time window (92–132 ms), mean ERP amplitudes were calculated for the four conditions (i.e., standard, syntactic irregularity, contour deviant, and double deviant) in the first and second sessions. In the first session, the average numbers of epochs were 361 (315–370), 98 (87–100), 362 (323–370), and 98 (83–100) epochs for the standard, contour deviant, syntactic irregularity, and double deviant (i.e., co-occurrence of syntactic irregularity and contour deviant) ERP waveforms, respectively. In the second session, the average number of epochs was 361 (315–370), 98 (86–100), 361 (319–370), and 98 (85–100) for the standard, contour deviant, syntactic irregularity, and double deviant ERP waveforms, respectively.

2.6. Statistical analysis

JASP 0.17.2 (JASP Team, 2023) was used to perform both classical (frequentist) and Bayesian analyses. Bayes Factors (BF_{-0} for one-tailed and BF_{10} for two-tailed) were calculated to assess the absence (null hypothesis) or presence (alternative hypothesis) of the difference. First, the difference in peak latency between the ERAN and MMN was tested using a paired *t*-test and a Bayesian paired *t*-test, using the detected negative peak in the 100–300 ms time window. Second, the presence of the ERAN and MMN was tested using a one-sample *t*-test (one-tailed) and a Bayesian one-sample *t*-test on the ERP amplitude of the ERAN and MMN intervals in each session, using ERAN and MMN amplitudes calculated from the defined time window (ERAN: 90–130 ms; MMN: 133–173 ms). The ERAN and MMN amplitudes were also compared between the first and second sessions using a paired *t*-test (two-tailed). Third, the interaction of irregularity and deviance factors was tested by a repeated measure two-way analysis of variance (ANOVA) with music-syntactic irregularity (with/without music-syntactic irregularity) and melodic-contour deviance (with/without melodic-contour deviance) separately in the first and second sessions, using the mean ERP amplitude extracted from the 92–132 ms time window. A Bayesian two-way ANOVA was also performed. Due to technical failure, two participants' fitness rating data were not available, and the remaining $N = 34$ data were submitted to the same two-way ANOVA in the first and second sessions. The difference ratings of the marginal means of music-syntactic irregularity ([syntactic irregularity + double deviant]/2) – ([standard + contour deviant]/2) and contour deviance ([contour deviant + double deviant]/2) – ([standard + syntactic irregularity]/2) were directly compared between two sessions. Also, a four-way ANOVA with factors of group (Group I/Group II), session, music-syntactic irregularity, and melodic-contour deviance on fitness ratings is reported in Supplementary Analysis S1. The significance levels were set at $\alpha = .05$ for frequentist hypothesis testing. The Cauchy distribution (scale parameter *r* of 0.707) was used as the prior distribution for δ in the Bayesian *t*-test, and the multivariate Cauchy distribution (fixed effect: scale parameter $r = 0.5$; random effect: scale parameter $r = 1$; covariates: scale parameter $r = 0.354$) was used as the prior distribution in the Bayesian two-way repeated measures ANOVA. For Bayesian hypothesis testing, a Bayes factor (BF_{01}) greater than 3 was considered moderate evidence for the null hypothesis (Schönbrodt and Wagenmakers, 2018). The stimulus materials and the data necessary to replicate the results are available at https://osf.io/49bys.

3. Results

3.1. Manipulation check

In the timbre detection task, the average of the mean reaction time of the timbre change detection task was 353 ms ($SD = 78$ ms), and the mean hit rate was 94.0% ($SD = 10.4\%$) across the ten blocks. Thus, the participants attended to the melodic sequence during the EEG recording.

In the first session, the mean fitness ratings (*SD*) were 6.2 (0.8), 3.0 (1.1), 5.9 (1.1), and 2.9 (1.2) for the standard, syntactic irregularity, contour deviant, and double deviant, respectively. The two-way ANOVA with music-syntactic irregularity and melodic-contour deviance on the fitness ratings revealed the significance of music-syntactic irregularity, $F(1, 33) = 161.92$, $p < .001$, $\eta_p^2 = 0.831$, $BF_{10} = 1.63 \times 10^{11}$, and melodic-contour deviance, $F(1, 33) = 6.86$, $p = .013$, $\eta_p^2 = 0.172$, $BF_{10} = 3.31$. However, the interaction was not significant, $F(1, 33) = 1.76$, $p = .194$, $\eta_p^2 = 0.051$, $BF_{10} = 0.58$. Supporting hypotheses H4a and H4b, the participants perceived that the music-syntactic irregularity and the melodic-contour deviant did not fit, reflecting the violation of schematic and dynamic expectations. However, as reported in the three-way ANOVA of Supplementary Analysis S1, the sense of deviance in Group II was weak.

In the second session, the mean fitness ratings (*SD*) were 6.0 (1.1), 2.8 (1.2), 6.3 (0.8), and 2.8 (1.1) for the standard, syntactic irregularity, contour deviant, and double deviant, respectively. The two-way ANOVA revealed the significance of music-syntactic irregularity, $F(1, 33) = 164.79.$, $p < .001$, $\eta_p^2 = 0.833$, $BF_{10} = 2.523 \times 10^{11}$, but not melodic-contour deviance, $F(1, 33) = 2.17$, $p = .151$, $\eta_p^2 = 0.062$, $BF_{10} = 0.57$, or the interaction, $F(1, 33) = 2.61$, $p = .116$, $\eta_p^2 = 0.073$, $BF_{10} = 0.73$. The participants still perceived the music-syntactic irregularity as the musical deviant but not the melodic-contour deviant, suggesting the effect of the difference in plasticity of expectations.

The decrease in fitness rating by syntactic irregularity was significantly lower in the second session ($M = -3.3$ $SD = 1.5$) than in the first session ($M = -3.1$, $SD = 1.4$), $t(33) = 2.31$, $p = .028$, $dz = 0.40$, $BF_{10} = 1.90$. In contrast, the decrease in fitness rating by contour deviance was significantly higher in the second session ($M = 0.1$, $SD = 0.6$) than in the first session ($M = -0.3$, $SD = 0.6$), $t(33) = -2.18$, $p = .036$, $dz = -0.37$, $BF_{10} = 1.47$. These results suggest that syntactic irregularity was perceived as more deviant in the second session and contour deviance was perceived as less deviant in the second session.

3.2. Differences in ERAN and MMN responses

Fig. 2 shows the grand average waveforms and scalp topographies of the ERAN and MMN. As shown in Fig. 2A, when collapsed across the sessions (i.e., the whole ten blocks), ERAN and MMN were observed around 100–200 ms in the difference waveforms calculated by subtracting the ERP waveforms for the notes without irregularity or deviance from those for the notes with irregularity or deviance in each session. Peak latencies were detected in the difference waveforms, and the ERAN and MMN time windows were determined separately (peak ±20 ms). The music-syntactic irregularity elicited the ERAN with a peak latency of 110 ms, and the ERAN amplitude ($M = -0.83$ μV, $SD = 0.84$) calculated in a period of 90–130 ms was significantly negative, $t\,(35) = -5.94$, $p < .001$, $d_z = -0.99$, $BF_{10} = 36638.636$. The melodic-contour deviant also elicited the MMN with a peak latency of 153 ms, and the MMN amplitude ($M = -0.20$ μV, $SD = 0.46$) calculated in a period of 133–173 ms was significantly negative, $t\,(35) = -2.65$, $p = .006$, $d_z = 0.44$, $BF_{10} = 7.11$. The peak latency of ERAN ($M = 145$ ms, $SD = 55$ ms) and MMN ($M = 193$ms, $SD = 64$ ms) was significantly different, $t\,(35) = -3.73$, $p < .001$, $d_z = -0.62$, $BF_{10} = 44.73$, and hypothesis H1 was not supported.

To examine the difference in plasticity, the mean ERP amplitudes calculated from the ERAN/MMN time windows were examined separately in the first and second sessions. Fig. 2B and C shows the results of the ERAN and MMN responses in the first and second sessions. In support of hypothesis H3a, the ERAN amplitudes were significant in both the first session ($M = -0.87$ μV, $SD = 0.84$), $t\,(35) = -6.21$, $p < .001$, $d_z = -1.04$, $BF_{-0} = 78579.34$, and second session ($M = -0.79$ μV, $SD = 0.98$), $t\,(35) = -4.80$, $p < .001$, $d_z = -0.80$, $BF_{-0} = 1525.12$. However, the MMN amplitude was significant only in the first session ($M = -0.46$ μV, $SD = 0.70$), $t\,(35) = -3.97$, $p < .001$, $d_z = -0.66$, $BF_{-0} = 164.19$, but not in the second session ($M = -0.06$ μV, $SD = 0.53$), $t\,(35) = -0.69$, $p = .247$, $d_z = -0.12$, $BF_{-0} = 0.33$, and hypothesis H3b was supported. These results suggest that the ERAN was still observed after the contour probability was reversed, but the MMN disappeared.

Furthermore, the ERAN and MMN amplitudes were directly compared between the first and second sessions. For the ERAN, amplitudes were not significantly different between the first and second sessions, $t\,(35) = -0.66$, $p = .516$, $dz = -0.11$, $BF_{10} = 0.22$. The Bayes factor provided moderate evidence for no difference. For the MMN, amplitudes differed significantly between the sessions, $t\,(35) = -2.65$, $p = .012$, $dz = -0.44$, $BF_{10} = 3.58$. The Bayes factor provided moderate evidence for a difference. These results were similar across the two groups; a Group × Session mixed ANOVA revealed no significant interaction effect (see Supplementary Material S2).

3.3. Relationship between schematic and dynamic expectations

Fig. 3 shows the grand average waveforms and scalp topographies of the ERPs elicited by the four types of melodies in each session. Because the latency of the ERAN and MMN partially overlapped, the interaction of each deviance factor was examined using a two-way ANOVA with music-syntactic irregularity and melodic-contour deviance on the ERP amplitudes over a period of 92–132 ms (peak latency 112 ms) in the first and second sessions, separately. As shown in the lower left panel of Fig. 3, the deviance-related ERP amplitude was greater for the double deviant ($M = -1.43$, $SD = 1.37$) than for the syntactic irregularity ($M = -0.83$, $SD = 0.87$) and contour deviant ($M = -0.37$, $SD = 0.69$) in the first session. However, the ERP amplitude was greater for the syntactic irregularity ($M = -0.85$, $SD = 1.06$) than for the double deviant ($M = -0.57$, $SD = 1.08$) and contour deviance ($M = -0.0008$, $SD = 0.67$) in the second session.

In the first session, syntactic irregularity, $F\,(1, 35) = 38.83$, $p < .001$, $\eta_p^2 = 0.526$, $BF_{10} = 44196.83$, and contour deviance were significant, $F\,(1, 35) = 19.73$, $p < .001$, $\eta_p^2 = 0.360$, $BF_{10} = 87.28$. However, the interaction was not significant, $F\,(1, 35) = 1.14$, $p = .292$, $\eta_p^2 = 0.032$, $BF_{10} = 0.43$, and hypothesis H2 was not supported. These results suggest a separate function of schematic and dynamic expectations. In the second session, only syntactic irregularity was significant, $F\,(1, 35) = 19.76$, $p < .001$, $\eta_p^2 = 0.361$, $BF_{10} = 307.75$, and insignificant results were shown for contour deviance, $F\,(1, 35) = 2.29$, $p = .139$, $\eta_p^2 = 0.061$, $BF_{10} = 0.46$, and the interaction, $F\,(1, 35) = 2.48$, $p = .124$, $\eta_p^2 = 0.066$, $BF_{10} = 0.83$. These results suggest that only the effect of music-syntactic irregularity was operative.

4. Discussion

The present study examined the difference in plasticity between schematic and dynamic expectations when each expectation simultaneously predicted the pitch class of the melody during music listening. The ERAN was elicited in both the first and second sessions, supporting hypothesis H3a. The MMN was elicited in the first session but disappeared in the second session, supporting hypothesis H3b. These results suggest that schematic expectations have less plasticity than dynamic expectations. Consistent with the ERP results, the syntactically irregular and contour deviant notes were both rated less fit in the first session, supporting hypotheses H4a and H4b, but in the second session, only the syntactically irregular notes were rated less fit. Although the ERAN and MMN occurred in a similar latency range (i.e., around 100–200 ms in Fig. 2A), the peak latency of the ERAN was shorter than that of the MMN, and hypothesis H1 was not supported. Moreover, when the syntactic irregularity and contour deviance co-occurred, the irregularity and deviance factors did not interact in either the first or second session, and hypothesis H2 was not supported. Thus, schematic and dynamic expectations functioned separately in this study.

The ERAN and MMN responses in the present study may be similar components in terms of latency and topographic distribution. However, we use the term "ERAN" because it had been established for the functional significance of this ERP component (Koelsch, 2009). Although the current ERAN response occurred earlier than that in previous studies (Koelsch et al., 2000, 2007), this may be due to stimulus characteristics. Koelsch and Jentschke (2010) reported that music-syntactic irregularity in melodies elicited an early anterior negativity (peaking at around 125 ms and named N125) and they considered it a subcomponent of the ERAN related to the processing of music-syntactic properties of melodies. Therefore, the current ERAN is not unusual in terms of peak latency.

Regarding the difference in the plasticity of expectations, hypotheses H3a and H3b were supported even when each expectation simultaneously predicted the same pitch class dimension. In his review, Koelsch (2009) proposed that the regularity representations underlying ERAN and MMN are different. ERAN is based on the long-term representation of music-syntactic knowledge, whereas MMN is based on regularities extracted from the current auditory context. Coinciding with this dissociation, the predictive processes reflected in the ERAN and MMN showed different plasticity in this study, with schematic expectation showing less plasticity than dynamic expectation. Koelsch and Jentschke (2008) demonstrated that the ERAN could be observed even when the syntactically irregular chord was presented with the same probability as the syntactically regular chord for 2 h. This finding suggests that the predictive model reflected in the ERAN cannot easily be deleted. In the present study, the ERAN was elicited in both the first and second sessions irrespective of the probability of irregular notes in the auditory context, suggesting less plasticity. Supportive results were also observed in the fitness ratings: the syntactic irregular was recognized as the deviance in both sessions, whereas the contour deviant was recognized as the deviance only in the first session. Therefore, the regularity representation of schematic expectations was stably maintained during music listening. In contrast, the MMN disappeared in the second session,

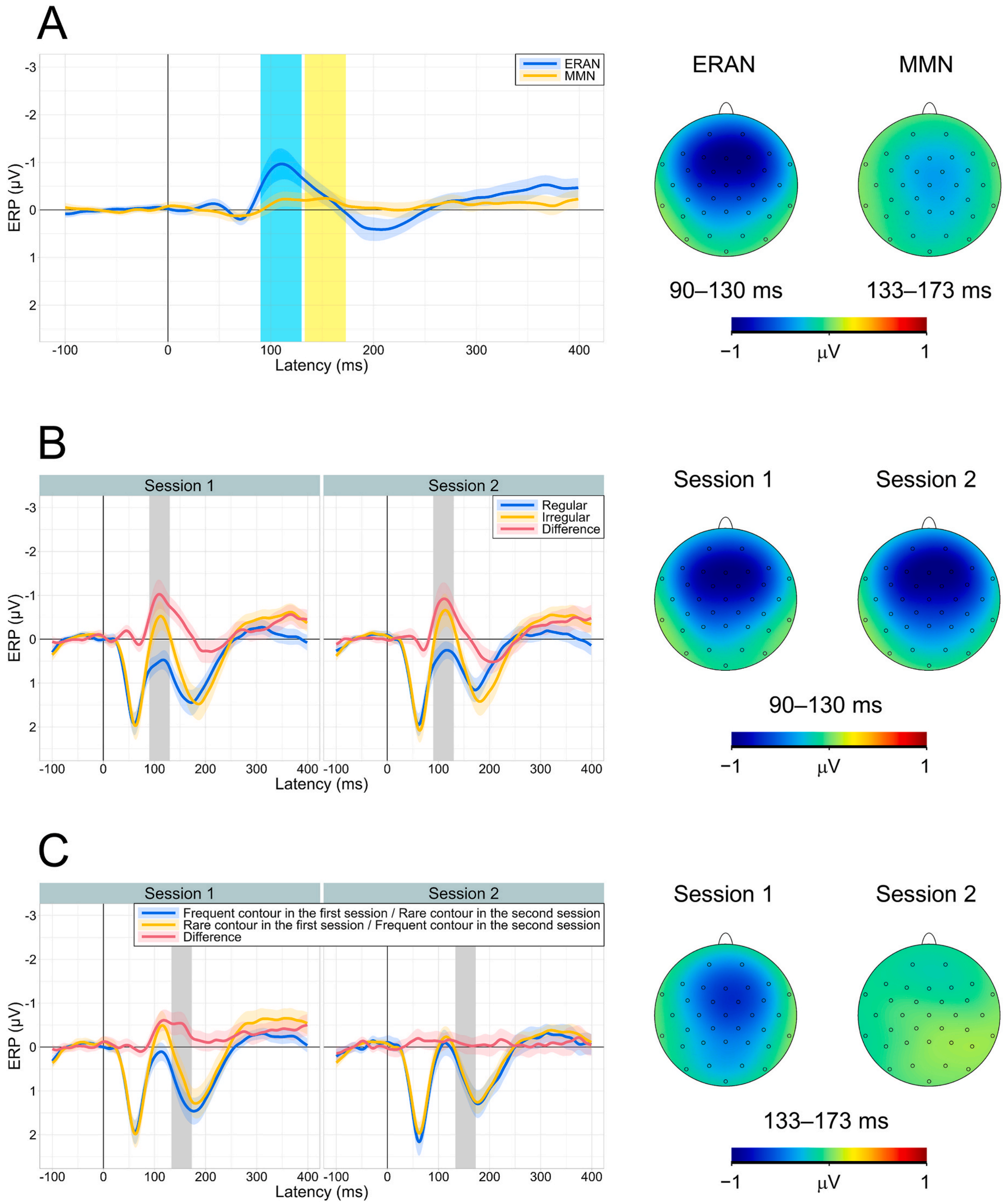

Fig. 2. ERP responses to music-syntactic irregularity and melodic-contour deviance.

Note. Panel A shows the ERAN and MMN responses collapsed across the two sessions (i.e., whole ten blocks). The difference waveforms were calculated by subtracting the ERP waveforms for the notes without irregularity or deviance from those for the notes with irregularity or deviance. Latency windows for amplitude analysis were determined based on these waveforms. The blue and yellow shaded areas indicate the ERAN (90–130 ms) and MMN (133–173 ms) time windows, respectively. Panels B and C show how the stimuli presented in the first session elicited the ERAN and MMN responses after the transitional probabilities were reversed. The ERAN in each session was calculated by subtracting the ERP waveforms elicited by music-syntactic regular notes from the ERP waveforms elicited by irregular notes, irrespective of their transitional probabilities. The MMN in the first session was calculated by subtracting the ERP waveforms elicited by frequent contour patterns from the ERP waveforms elicited by rare contour patterns, irrespective of their syntactic regularity. Conversely, the MMN in the second session was calculated by subtracting the ERP waveforms elicited by rare contour patterns (i.e., frequent contour patterns in the first session) from the ERP waveforms elicited by frequent contour patterns (i.e., rare contour patterns in the first session). The gray areas indicate the intervals determined in Panel A. The ERP waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) are shown with 95% confidence intervals.

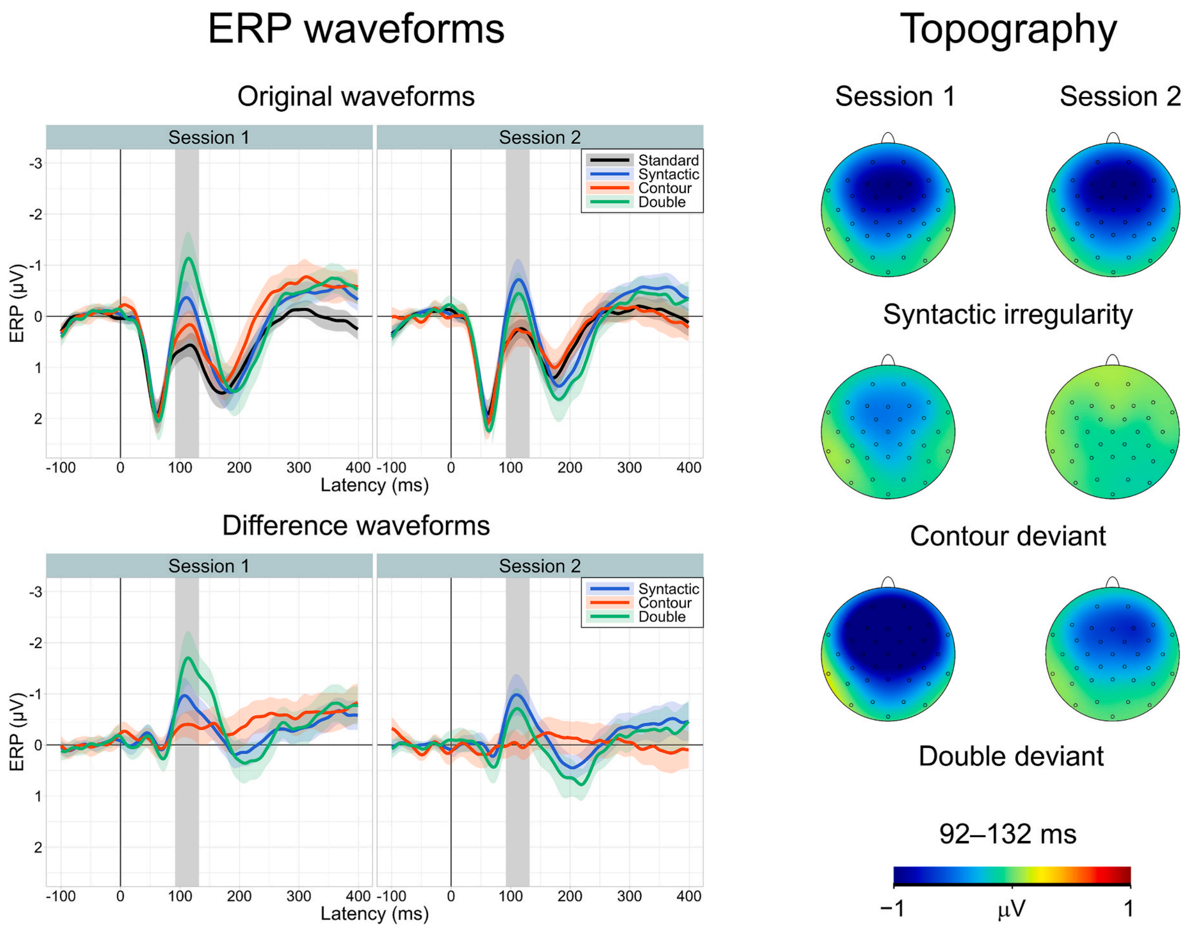

Fig. 3. ERP responses to the four types of melodies and deviance-related difference waveforms.
Note. The upper left panel shows the grand average ERP waveforms (means of the five frontal electrodes: F7, F3, Fz, F4, and F8) with 95% confidence intervals for the final note of each melody. The lower left panel shows the difference waveforms of three deviant melodies (deviants − standard). The gray area indicates the time window for the statistical analysis (92–132 ms) of the two-way ANOVA, which was determined based on the peak latency extracted in the average waveform of the ERAN and MMN grand mean waveforms. The right panel shows the topographic distributions of the deviance-related ERPs.

suggesting a shorter-term plasticity of the regularity representation in dynamic expectations. However, the sustainability of ERAN may be due to the out-of-key note, which has a higher salience, and future research should replicate the current results using the in-key music-syntactic irregular notes.

The short-term plasticity of deviance detection, particularly as reflected in the auditory MMN, has been examined both empirically and computationally (Lieder et al., 2013; Sussman & Winkler, 2001; reviewed in Kobayashi et al., 2024). It seems plausible that the predictive system that updates a generative model point by point in the short term, as in general auditory perception, also operates in music perception. Meanwhile, the lower plasticity observed in schematic expectations provides important insight into music perception. Music perception is based in part on schematic knowledge of musical structures and the regularities that constitute music (Bharucha and Krumhansl, 1983; DeWitt and Samuel, 1990; Koelsch, 2009; Sears et al., 2019), and these representations need to remain unchanged over time. If knowledge of musical syntax changed immediately during music listening, listeners would not be able to perceive the patterned chord progressions that appeared throughout the song, and the processing cost of music listening would increase.

Predictive coding frameworks emphasize that generative models underlying predictions about the world are dynamically updated (Garrido et al., 2009; Lieder et al., 2013). However, stable predictive models are also necessary for music perception. Sussman and Winkler (2001), who examined dynamic auditory deviance detection systems, noted that there may be a bias toward maintaining the current model until sufficient information about model updates is obtained, as the presence or absence of MMN does not change immediately with changes in the context when the deviant occurs. It is reasonable to assume that the generative models of harmonically and rhythmically structured music are not only updated on the basis of short-term plasticity but also maintained by a bias that preserves the current model of musical schema. This framework is consistent with previous research suggesting that music perception involves both bottom-up sensory processing and cognitive schema-based processing (Bigand et al., 1996, 2003).

Although higher plasticity was observed for dynamic expectations than for schematic expectations, the rare contour in the second session did not elicit the MMN. As shown in Fig. 2C, because the MMN time window of the difference waveform is around 0 μV, its sign-reversed waveform (i.e., contour deviance – contour standard in the second session) is also around 0 μV. Given the nature of dynamic expectations, it would be natural for the MMN not only to disappear with the regularities of the first session but also to appear with the regularities of the second session. However, this was not the case. This may suggest that the modulation of dynamic expectations takes some time and is not fully updated immediately, even though it is more plastic than schematic expectations. Once established, the predictive model of dynamic expectations is not updated as quickly, and the update may not occur until sufficient information is obtained. Daikoku et al. (2017) reported that correcting statistical learning takes longer than acquiring a new statistical regularity by comparing ERP responses before and after reversing the transition probabilities of tones. Therefore, updating dynamic expectations takes longer than establishing dynamic expectations for the first time. This is similar to the maintenance bias observed by Sussman and Winkler (2001).

Another possibility is that the formation of contour memory through statistical learning in the first session may generate a third type of expectations, memory-based *veridical* expectations (Bharucha and Todd, 1989), in the second session. In the subjective ratings, the contour deviant was rated as less fit only in the first session, and in the second session, the contour deviant (rare contour patterns in the first session) was rated as more fit than the standard (frequent contour patterns in the first session). Thus, veridical expectations from contour memory acquired in the first session may have competed with dynamic expectations, resulting in a lack of prediction error.

The absence of MMN elicitation by the contour deviant in the second session suggests that dynamic expectations are not fully and immediately updated by the ongoing auditory context (i.e., change of stimulus probabilities). Nevertheless, the facts that the contour deviant of the first session did not elicit the MMN in the second session and that the syntactic irregularity of the first session continued to elicit the ERAN in the second session provide evidence for a higher plasticity of predictive processes in dynamic expectations than in schematic expectations.

The interaction of music-syntactic irregularity and melodic-contour deviance was not observed in either session. The discrepancy between the present study and Ishida and Nittono (2024) may be due to the presentation probability in the context. In their study, a steep probability gradient was set for the contour transition patterns (90:10) compared to the present experiment (80:20). There are two possibilities for the effect of probability on the interaction. First, as reported by Ishida and Nittono, antagonism may occur in the single-deviant condition. However, this antagonism was weaker in the present study than in their study because the prediction error of the syntactic irregular condition was less inhibited by the non-violation of the contour deviance in this study ($p = .80$) than in the previous study ($p = .90$). As evidence, the ERAN amplitude in the syntactic irregular condition was numerically larger in the present study ($M = -0.83$ μV) than in the previous study ($M = -0.43$ μV). Therefore, the ERAN elicitation in the single-deviant condition may not have been as different as that in the double-deviant condition, resulting in the absence of multiplication.

Second, the multiplicative effect found by Ishida and Nittono (2024) may be due to the higher precision in the steep probability gradient, while the lack of multiplication may be due to the lower precision in the loose probability gradient in the present study. Entropy based on the probability of an event can affect statistical learning (Agres et al., 2018; Okano et al., 2021). In the MEG study, Okano et al. showed that a significant reduction in P1m was observed for lower device probability ($p = .10$) but not for higher deviance probability ($p = .33$). Moreover, the predictive coding framework proposes that prediction error is weighted by precision (Feldman and Friston, 2010; Friston and Kiebel, 2009), and Quiroga-Martinez et al. (2019) empirically showed that the MMN as a neural prediction error was attenuated when predictability was lower (i.e., high entropy) compared to when it was higher. The higher precision in Ishida and Nittono (2024) may have facilitated the ERAN elicitation in the double-deviant condition and, conversely, may not have facilitated the ERP elicitation in this study. Therefore, the multiplicative effect was absent. In any case, this suggests that ERAN elicitation is somewhat independent of short-term exposure but is indirectly influenced by the probabilistic information in the current auditory context. Koelsch and Jentschke (2008), who observed the ERAN even during the 2 h of exposure, also reported that the ERAN amplitude decreased linearly with continued exposure. Considering this, the interpretation that ERAN is slightly influenced by probability in the current musical context may be valid.

In conclusion, this study revealed less plasticity of schematic expectations than of dynamic expectations even when expectations concurrently predicted the same pitch class dimension. While previous studies have examined how multiple expectations work interactively in music perception (Ishida and Nittono, 2022, 2024; Tillmann and Bigand, 2010), the present study empirically demonstrates that different types of expectations (i.e., schematic and dynamic) show different plasticity by changing the transitional probabilities in the middle of the experiment. Considering natural music listening, it is reasonable that dynamic expectations, which are dynamically updated through shorter-term plasticity, and schematic expectations, which are cautiously updated through longer-term plasticity, function separately in music perception. Although an interaction between the two expectations was not observed in this study, the potential influence of entropy or precision through contextual probability on the interaction of expectations remains an area for future investigation.

CRediT authorship contribution statement

Kai Ishida: Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Hiroshi Nittono:** Writing – review & editing, Validation, Supervision, Software, Resources, Methodology, Formal analysis, Conceptualization.

Declaration of competing 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.

Acknowledgement

This work was supported by JSPS KAKENHI JP22KJ2199.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2024.109012.

Data availability

The sound materials used and datasets analyzed for the present paper are available at https://osf.io/49bys.

References

Agres, K., Abdallah, S., Pearce, M., 2018. Information-theoretic properties of auditory sequences dynamically influence expectation and memory. Cognit. Sci. 42, 43–76. https://doi.org/10.1111/cogs.12477.
Bharucha, J., Krumhansl, C.L., 1983. The representation of harmonic structure in music: hierarchies of stability as a function of context. Cognition 13, 63–102. https://doi.org/10.1016/0010-0277(83)90003-3.
Bharucha, J.J., Todd, P.M., 1989. Modeling the perception of tonal structure with neural nets. Comput. Music J. 13, 44–53. https://doi.org/10.2307/3679552.
Bigand, E., Parncutt, R., Lerdahl, F., 1996. Perception of musical tension in short chord sequences: the influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Percept. Psychophys. 58, 125–141. https://link.springer.com/article/10.3758/BF03205482.
Bigand, E., Poulin, B., Tillmann, B., Madurell, F., D'Adamo, D.A., 2003. Sensory versus cognitive components in harmonic priming. J. Exp. Psychol. Human 29, 159–171. https://doi.org/10.1037/0096-1523.29.1.159.
Brattico, E., Tervaniemi, M., Näätänen, R., Peretz, I., 2006. Musical scale properties are automatically processed in the human auditory cortex. Brain Res. 1117, 162–174. https://doi.org/10.1016/j.brainres.2006.08.023.
Daikoku, T., Yatomi, Y., Yumoto, M., 2017. Statistical learning of an auditory sequence and reorganization of acquired knowledge: a time course of word segmentation and ordering. Neuropsychologia 95, 1–10. https://doi.org/10.1016/j.neuropsychologia.2016.12.006.
Dewitt, L.A., Samuel, A.G., 1990. The role of knowledge-based expectations in music perception: evidence from musical restoration. J. Exp. Psychol. Gen. 119, 123–144. https://doi.org/10.1037//0096-3445.119.2.123.
Faul, F., Erdfelder, E., Lang, A.G., Buchner, A., 2007. G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191. https://doi.org/10.3758/BF03193146.

Feldman, H., Friston, K.J., 2010. Attention, uncertainty, and free-energy. Front. Hum. Neurosci. 4, 215. https://doi.org/10.3389/fnhum.2010.00215.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B, Biol. Sci. 364, 1211–1221. https://doi.org/10.1098/rstb.2008.0300.

Garrido, M.I., Kilner, J.M., Stephan, K.E., Friston, K.J., 2009. The mismatch negativity: a review of underlying mechanisms. Clin. Neurophysiol. 120, 453–463. https://doi.org/10.1016/j.clinph.2008.11.029.

Guo, S., Koelsch, S., 2016. Effects of veridical expectations on syntax processing in music: event-related potential evidence. Sci. Rep. 6 (1), 19064. https://doi.org/10.1038/srep19064.

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

Ishida, K., Nittono, H., 2024. Relationship between schematic and dynamic expectations of melodic patterns in music perception. Int. J. Psychophysiol. 196, 112292. https://doi.org/10.1016/j.ijpsycho.2023.112292.

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

Kobayashi, K., Shiba, Y., Honda, S., Nakajima, S., Fujii, S., Mimura, M., Noda, Y., 2024. Short-term effect of auditory stimulation on neural activities: a scoping review of longitudinal electroencephalography and magnetoencephalography studies. Brain Sci. 14, 131. https://doi.org/10.3390/brainsci14020131.

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

Koelsch, S., Gunter, T.C., Schro, E., Tervaniemi, M., Sammler, D., Friederici, A.D., 2001. Differentiating ERAN and MMN: an ERP study. Neuroreport 12, 1385–1389. https://doi.org/10.1097/00001756-200105250-00019.

Koelsch, S., Gunter, T., Friederici, A.D., Schröger, E., 2000. Brain indices of music processing: "Nonmusicians'' are musical. J. Cognit. Neurosci. 12, 520–541. https://doi.org/10.1162/089892900562183.

Koelsch, S., Jentschke, S., 2008. Short-term effects of processing musical syntax: an ERP study. Brain Res. 1212, 55–62. https://doi.org/10.1016/j.brainres.2007.10.078.

Koelsch, S., Jentschke, S., 2010. Differences in electric brain responses to melodies and chords. J. Cognit. Neurosci. 22 (10), 2251–2262. https://doi.org/10.1162/jocn.2009.21338.

Koelsch, S., Jentschke, S., Sammler, D., Mietchen, D., 2007. Untangling syntactic and sensory processing: an ERP study of music perception. Psychophysiology 44, 476–490. https://doi.org/10.1111/j.1469-8986.2007.00517.x.

Koelsch, S., Rohrmeier, M., Torrecuso, R., Jentschke, S., 2013. Processing of hierarchical syntactic structure in music. Proc. Natl. Acad. Sci. U.S.A. 110, 15443–15448. https://doi.org/10.1073/pnas.1300272110.

Koelsch, S., Sammler, D., 2008. Cognitive components of regularity processing in the auditory domain. PLoS One 3, e2650. https://doi.org/10.1371/journal.pone.0002650.

Koelsch, S., Vuust, P., Friston, K., 2019. Predictive processes and the peculiar case of music. Trends Cognit. Sci. 23, 63–77. https://doi.org/10.1016/j.tics.2018.10.006.

Leino, S., Brattico, E., Tervaniemi, M., Vuust, P., 2007. Representation of harmony rules in the human brain: further evidence from event-related potentials. Brain Res. 1142, 169–177. https://doi.org/10.1016/j.brainres.2007.01.049.

Lieder, F., Daunizeau, J., Garrido, M.I., Friston, K.J., Stephan, K.E., 2013. Modelling trial-by-trial changes in the mismatch negativity. PLoS Comput. Biol. 9, e1002911. https://doi.org/10.1371/journal.pcbi.1002911.

Miranda, R.A., Ullman, M.T., 2007. Double dissociation between rules and memory in music: an event-related potential study. Neuroimage 38, 331–345. https://doi.org/10.1016/j.neuroimage.2007.07.034.

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

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

Okubo, M., Suzuki, H., Nicholls, M.E.R., 2014. A Japanese version of the FLANDERS handedness questionnaire. Shinrigaku Kenkyu 85, 474–481. https://doi.org/10.4992/jjpsy.85.13235.

Quiroga-Martinez, D.R., Hansen, N.C., Højlund, A., Pearce, M.T., Brattico, E., Vuust, P., 2019. Reduced prediction error responses in high-as compared to low-uncertainty musical contexts. Cortex 120, 181–200. https://doi.org/10.1016/j.cortex.2019.06.010.

Rohrmeier, M.A., Koelsch, S., 2012. Predictive information processing in music cognition. A critical review. Int. J. Psychophysiol. 83, 164–175. https://doi.org/10.1016/j.ijpsycho.2011.12.010.

Schönbrodt, F.D., Wagenmakers, E.J., 2018. Bayes factor design analysis: planning for compelling evidence. Psychon. Bull. Rev. 25, 128–142. https://doi.org/10.3758/s13423-017-1230-y.

Sears, D.R.W., Pearce, M.T., Spitzer, J., Caplin, W.E., McAdams, S., 2019. Expectations for tonal cadences: sensory and cognitive priming effects. Q. J. Exp. Psychol. 72, 1422–1438. https://doi.org/10.1177/1747021818814472.

Sussman, E.S., Chen, S., Sussman-Fort, J., Dinces, E., 2014. The five myths of MMN: redefining how to use MMN in basic and clinical research. Brain Topogr. 27, 553–564. https://doi.org/10.1007/s10548-013-0326-6.

Sussman, E., Winkler, I., 2001. Dynamic sensory updating in the auditory system. Cogn. Brain Res. 12 (3), 431–439. https://doi.org/10.1016/S0926-6410(01)00067-2.

Tillmann, B., Bigand, E., 2010. Musical structure processing after repeated listening: schematic expectations resist veridical expectations. Music. Sci. 14 (2_Suppl. l), 33–47. https://doi.org/10.1177/10298649100140S204.

Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., Tervaniemi, M., 2011. The preattentive processing of major vs. minor chords in the human brain: an event-related potential study. Neurosci. Lett. 487, 406–410. https://doi.org/10.1016/j.neulet.2010.10.066.

Vuust, P., Heggli, O.A., Friston, K.J., Kringelbach, M.L., 2022. Music in the brain. Nat. Rev. Neurosci. 23, 287–305. https://doi.org/10.1038/s41583-022-00578-5.

Vuust, P., Ostergaard, L., Pallesen, K.J., Bailey, C., Roepstorff, A., 2009. Predictive coding of music - brain responses to rhythmic incongruity. Cortex 45, 80–92. https://doi.org/10.1016/j.cortex.2008.05.014.