

Pitch-class distribution modulates the statistical learning of atonal chord sequences



Tatsuya Daikoku, Yutaka Yatomi, Masato Yumoto*

Department of Clinical Laboratory, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Article history:

Received 12 May 2015

Revised 22 March 2016

Accepted 28 June 2016

Keywords:

Statistical learning

Pitch class

Musical chord sequence

Markov process

Magnetoencephalography

ABSTRACT

The present study investigated whether neural responses could demonstrate the statistical learning of chord sequences and how the perception underlying a pitch class can affect the statistical learning of chord sequences. Neuromagnetic responses to two chord sequences of augmented triads that were presented every 0.5 s were recorded from fourteen right-handed participants. One sequence was a series of 360 chord triplets, each of which consisted of three chords in the same pitch class (clustered pitch-classes sequences). The other sequence was a series of 360 chord triplets, each of which consisted of three chords in different pitch classes (dispersed pitch-classes sequences). The order of the triplets was constrained by a first-order Markov stochastic model such that a forthcoming triplet was statistically defined by the most recent triplet (80% for one; 20% for the other two). We performed a repeated-measures ANOVA with the peak amplitude and latency of the P1m, N1m and P2m. In the clustered pitch-classes sequences, the P1m responses to the triplets that appeared with higher transitional probability were significantly reduced compared with those with lower transitional probability, whereas no significant result was detected in the dispersed pitch-classes sequences. Neuromagnetic significance was concordant with the results of familiarity interviews conducted after each learning session. The P1m response is a useful index for the statistical learning of chord sequences. Domain-specific perception based on the pitch class may facilitate the domain-general statistical learning of chord sequences.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

How do we learn the structure embedded in auditory sequences such as language and music? Although both language and music have syntax (Lerdahl & Jackendoff, 1983), the details of the syntactic structure are different (Keiler, 1978). In recent studies, much attention has been given to domain-general learning versus domain-specific learning. Domain-specific learning suggests that each type of knowledge for handling tasks in a particular domain is specialized (Hauser, Chomsky, & Fitch, 2002). Domain-specific processing in music may be evidenced by the selective impairments in music recognition abilities (Peretz & Coltheart, 2003). Individuals with congenital amusia can no longer recognize melodies without words, while they are normal in recognizing spoken words (Ayotte, Peretz, & Hyde, 2002). In contrast, domain-general learning indicates that almost all knowledge for handling tasks in all domains can be generalized. In other words, a common learning

mechanism may be applied to the auditory stream in different domains (François & Schön, 2014; Frensch & Buchner, 1999; Schön & François, 2011; Schön et al., 2008). For example, novices such as infants should recognize and learn their native language without specific linguistic knowledge, such as words, syntax, and grammar, implying that domain-general learning is a basic mechanism innate to humans. One of the domain-general learning mechanisms that has been suggested in recent studies is statistical learning. In previous studies, learners could acquire the statistical regularities that are involved in transitional probabilities and are embedded in speech tone (François & Schön, 2014; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Saffran, Loman, & Robertson, 2001) or chord sequences (Jonaitis & Saffran, 2009; Smith & Melara, 1990), with no specific knowledge of language and music domains: phonemes, syllables, syntax, and pitch classes (Hauser et al., 2002; Jackendoff & Lerdahl, 2006). Statistical learning has been demonstrated in not only behavioural but also neurophysiological results (Abla, Katahira, & Okanoya, 2008; Daikoku, Yatomi, & Yumoto, 2014, 2015; François & Schön, 2014; Furl et al., 2011; Koelsch, Busch, Jentschke, & Rohrmeier, 2016; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012; Schön & François, 2011).

* Corresponding author at: Department of Clinical Laboratory, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan.

E-mail address: yumoto-tyk@umin.ac.jp (M. Yumoto).

According to previous research on music perception, pitch perception may partially depend on the domain-specific processing component called pitch class, which is a set of pitches that are separated by one or more octaves (Drobisch, 1855; Fig. 1). All pitches that belong to the same pitch class have the same name (Whittall, 2008). For example, pitch class A consists of all pitches of A's in octaves (i.e., 440×2^n Hz; 55, 110, 220, 440, 880 Hz...). The pitch class has an important role in introducing domain-specific regularities into music. However, music learning mechanisms cannot be fully explained by domain-specific rules in music because musical principles vary cross-culturally (Castellano, Bharucha, & Krumhansl, 1984). Thus, the general hypothesis suggests that musical structures may be acquired by both the domain-general and domain-specific learning mechanisms (Jonaitis & Saffran, 2009; Tillmann, Bharucha, & Bigand, 2000).

Using violation paradigms, previous studies have investigated how our brain processes not only language but also musical syntax. In previous studies on semantic processing, the event-related potential (ERP) N400 was elicited by the word that semantically mismatched the preceding context in a linguistic sequence (Kutas & Hillyard, 1980), peaking at approximately 400 ms after the onset of the stimulus. The early left anterior negativity (ELAN) between 150 and 250 ms in latency could reflect early syntactic processes in language learning. Friederici et al. reported that the ELAN was elicited by a syntactic violation of phrase structures (Friederici, Pfeifer, & Hahne, 1993), even without attention to the stimuli (Friederici, 1995). The P600 component could reflect a late syntactic reanalysis as a marker for syntactic integration difficulty (Kaan, Harris, Gibson, & Holcomb, 2000).

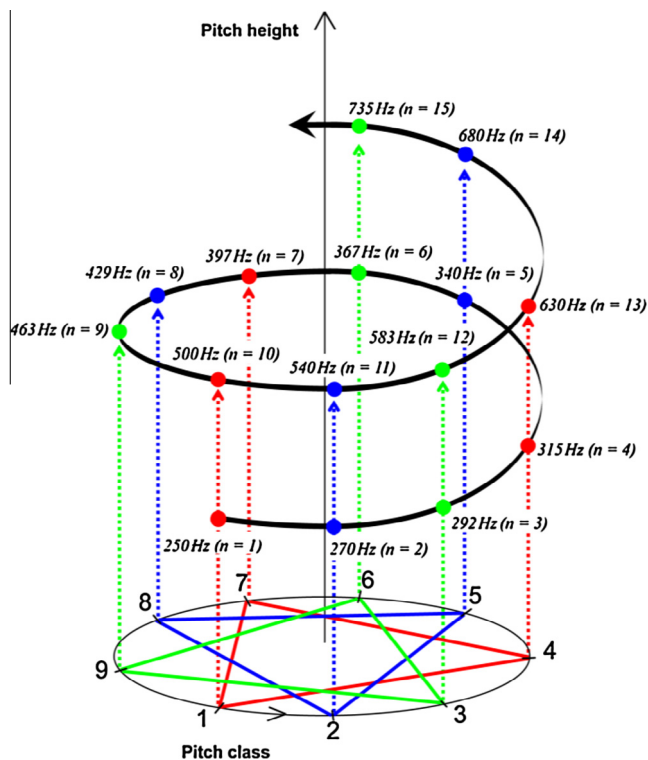


Fig. 1. The pitch class-pitch height space was based on the nine-tone equal temperament ($F_0 = 250 \times 2^{(n-1)/9}$ Hz, $n = 1-15$) in the present study. The circle at the bottom of the spiral helix is the pitch class dimension that is based on the nine-tone equal temperament. The dots where the arrow intersects the spiral are pitches in the same pitch class. For example, pitch class 1 in the circle at the bottom of the spiral consists of 250 Hz and 500 Hz, which are pitches that are one octave apart. Each colour represents three distinct pitch classes.

The ERP components could also reflect the syntactic processing of music. The N400 could reflect transgressed endings of chord sequences in music (James, Cereghetti, Rouillet, & Oechslin, 2015). The early right anterior negativity (ERAN), which was elicited by deviant chords embedded in chord progression in the same latency range of the ELAN, could reflect syntactic processes in music recognition (Kim, Kim, & Chung, 2011; Koelsch, Gunter, Friederici, & Schroger, 2000; Koelsch, 2009). Repeated exposure to deviant chords in chord sequences did not influence the ERAN magnitude (Guo & Koelsch, 2015), suggesting that short-term experience could not modulate the early evoked responses such as the ERAN. A previous study demonstrated that conditional probabilities in Western music corpus were reflected in the ERAN even in non-musicians, suggesting that the long-term musical knowledge is acquired by statistical learning (Kim et al., 2011).

Novices, such as infants, can learn musical syntax such as melody and chord progression without a priori knowledge of the theory of music, at least at an earlier stage of learning (Trehub & Hannon, 2006). In the present study, we tried to emulate this environment such that the participants were exposed to a novel syntactic structure of chord sequences that had never been experienced by adopting a repetition of augmented triads in a nine-tone equal temperament system. The augmented triad can occasionally appear in tonal music. However, a consecutive series of augmented triads with changing scales in a nine-tone equal temperament is no longer tonal and never occurs in Western music, unless detuned instruments are prepared. Moreover, the syntax (i.e., the order of chords) was under the control of an original rule independent of a Western music corpus, although a constant frequency interval between the adjacent tones in augmented triads (i.e., a trisection of the octave) might be considered to be borrowed from the Western music system. Accordingly, participants could no longer depend on the long-term knowledge of Western harmonic progression, and the statistical learning of a novel syntactic structure of chord sequences could be studied beyond cross-cultural differences in musical principles. Although previous studies revealed that statistical learning of chord sequences could be reflected in behavioural responses (Jonaitis & Saffran, 2009; Smith & Melara, 1990), to the best of our knowledge, no study has investigated how the statistical learning of novel syntactic structures of atonal chord sequences can be reflected in neurophysiological responses.

According to recent studies, the statistical learning of single tone sequences with higher-order structure introduced by a Markov stochastic model or a word segmentation paradigm is reflected in neuromagnetic responses, such as P1m, N1m, and P2m (the magnetic counterpart of the P1, N1, and P2 potentials) (Daikoku et al., 2014, 2015; Furl et al., 2011; Koelsch et al., 2016; Paraskevopoulos et al., 2012). In the studies using a word segmentation paradigm in which sequences were comprised of tone words, neural responses to the final tones within the words were significantly reduced compared to the initial tones, due to the contrast between high across-word variability and low within-word variability. In the studies using Markov chains, the neural responses to tones that appeared with higher transitional probability were significantly reduced compared to those with lower transitional probability. The Markov chain, which was first reported by Markov (1971), is a mathematical system in which the probability of the forthcoming state is statistically defined by the most recent state. The word segmentation paradigm is a specific form of the Markov chain. In several fields of study, including natural language processing (Poon & Domingos, 2007, 2008; Singla & Domingos, 2006), music perception and statistical learning (Richardson & Domingos, 2006), the Markov chain has been used as a model for the artificial grammar of language and music. The use of the Markov chains embedded in auditory sequences allows us to verify

statistical learning mechanisms in language and music acquisition. In the present study, we used the Markov stochastic model to investigate how domain-specific pitch-class perception and domain-general statistical learning of the chord sequences were reflected in neuromagnetic responses.

As described earlier in this paper, musical structures may be acquired by both the domain-general and domain-specific learning mechanisms (Jonaitis & Saffran, 2009; Tillmann et al., 2000). The present study aimed to evaluate the hypothesis that the domain-specific perception underlying the pitch class could affect the domain-general statistical learning of chord sequences.

2. Methods

2.1. Participants

Participants were fourteen right-handed (Edinburgh handedness questionnaires; laterality quotient ranged from 52 to 100; Oldfield, 1971) healthy Japanese participants (7 males, 7 females; age range: 22–50 years) who had no history of neurological or audiological disorders. None of the participants were professional musicians, and no participants had absolute pitch, according to self-report. This study was approved by the Ethics Committee of The University of Tokyo. All participants were informed of the purpose, safety and protection of personal data, and provided written informed consent to participate in this study.

2.2. Stimuli

2.2.1. Chords

Using pure tones with equal amplitude and frequencies in a nine-tone equal temperament ($F_0 = 250 \times 2^{(n-1)/9}$ Hz, $n = 1-15$) (Fig. 1), we generated nine chords of augmented triads consisting of two major thirds: Chord 1 (250, 315, 397 Hz), Chord 2 (315, 397, 500 Hz), Chord 3 (397, 500, 630 Hz), Chord 4 (270, 340, 429 Hz), Chord 5 (340, 429, 540 Hz), Chord 6 (429, 540, 680 Hz), Chord 7 (292, 367, 463 Hz), Chord 8 (367, 463, 583 Hz), and Chord 9 (463, 583, 735 Hz). The duration of each chord was 450 ms, which included a Gaussian rise and fall of 10 and 200 ms. The maximum slope of the Gaussian rise was double that of the 10-ms linear rise at 7 ms after tone onset.

2.2.2. Chord groupings

The nine chords could be categorized into three groups based on pitch-classes (see Table 1a). Furthermore, the nine chords could also be categorized into three groups across pitch-classes (see Table 1b).

2.2.3. Sequences

We devised two sequences with 1080 chords, which were presented with a constant stimulus onset asynchrony (SOA) of 0.5 s. The “clustered pitch-classes sequences” was a concatenation of chord triplets (Fig. 2a and c) that belonged to either group of A, B or C (Table 1a). The “dispersed pitch-classes sequences” was a

concatenation of chord triplets (Fig. 2b and d) that belonged to either group of D, E or F (Table 1b). In either sequence, the order of the “three chords” within each triplet was randomised, whereas the order of the triplet group was constrained based on a first-order Markov chain such that the forthcoming group was statistically defined by the most recent group: the recursive transitions (i.e., repetition of a group) with 14% probability, and the clockwise (e.g., $A \rightarrow B$ and $D \rightarrow E$) and counterclockwise transitions (e.g., $B \rightarrow A$ and $E \rightarrow D$) with 80% and 6% probabilities, respectively (Fig. 2). In other words, between-group ordering was defined by a Markov chain, while within-group ordering was randomised. In the chord presentations, a one-second silent period was pseudo-randomly inserted (i.e., SOA 1.5 s) somewhere in every set of 40 successive chords. Participants were instructed to raise their right hand at every silent period to confirm that they were paying attention to the chord sequences.

2.3. Measurement

We recorded magnetoencephalographic (MEG) signals from participants while they listened to two 1080-chord sequences. The order of the two chord sequences was counterbalanced across participants. Auditory stimuli were sequenced with the STIM2 system (Compumedics Neuroscan, El Paso, TX, USA) and were binaurally delivered to participants' ears at 80 dB SPL through ER-3A earphones (Etymotic Research, Elk Grove Village, IL, USA). The conduction delay from subminiature speakers in the earphones to the ear drums (1.6 ms) was corrected in the analysis. MEG signals were recorded in a magnetically shielded room, using a 306-channel neuromagnetometer system (Elekta Neuromag Oy, Helsinki, Finland), which has 204 planar first-order gradiometers and 102 magnetometers at 102 measuring sites on a helmet-shaped surface that covers the entire scalp. Auditory stimulus-triggered epochs were filtered online with a 0.1–200 Hz band-pass filter and were then recorded at a sampling rate of 600 Hz.

2.4. Behavioural tests

After the MEG measurement for each chord sequence, participants conducted a familiarity test in which they were presented with thirty series of 15 chords. They reported whether each 15-chord series sounded familiar or not. The thirty series, which consisted of three types including ten series each, were randomly ordered (Fig. 2).

In the first types of the clustered (Markov triplet series; Fig. 2e) and dispersed (Fig. 2h) pitch-classes sessions, the chord ordering was regulated by the same statistical constraints as the clustered and dispersed pitch-classes sequences used in the MEG measurement session, respectively. In other word, the order of the three chords within each triplet from a Group was randomised, while the order of the triplet groups was constrained based on the same Markov chain as the clustered and dispersed pitch-classes sequences. In the second types of the clustered (random triplet series; Fig. 2f) and dispersed (Fig. 2i) pitch-classes sessions, the

Table 1
Chord groups in the clustered pitch-classes (a) and dispersed pitch-classes sequences (b).

a. Clustered pitch-classes sequences			
Group A	Chord 1 (250, 315, 397 Hz)	Chord 2 (315, 397, 500 Hz)	Chord 3 (397, 500, 630 Hz)
Group B	Chord 4 (270, 340, 429 Hz)	Chord 5 (340, 429, 540 Hz)	Chord 6 (429, 540, 680 Hz)
Group C	Chord 7 (292, 367, 463 Hz)	Chord 8 (367, 463, 583 Hz)	Chord 9 (463, 583, 735 Hz)
b. Dispersed pitch-classes sequences			
Group D	Chord 1 (250, 315, 397 Hz)	Chord 5 (340, 429, 540 Hz)	Chord 9 (463, 583, 735 Hz)
Group E	Chord 2 (315, 397, 500 Hz)	Chord 6 (429, 540, 680 Hz)	Chord 7 (292, 367, 463 Hz)
Group F	Chord 3 (397, 500, 630 Hz)	Chord 4 (270, 340, 429 Hz)	Chord 8 (367, 463, 583 Hz)

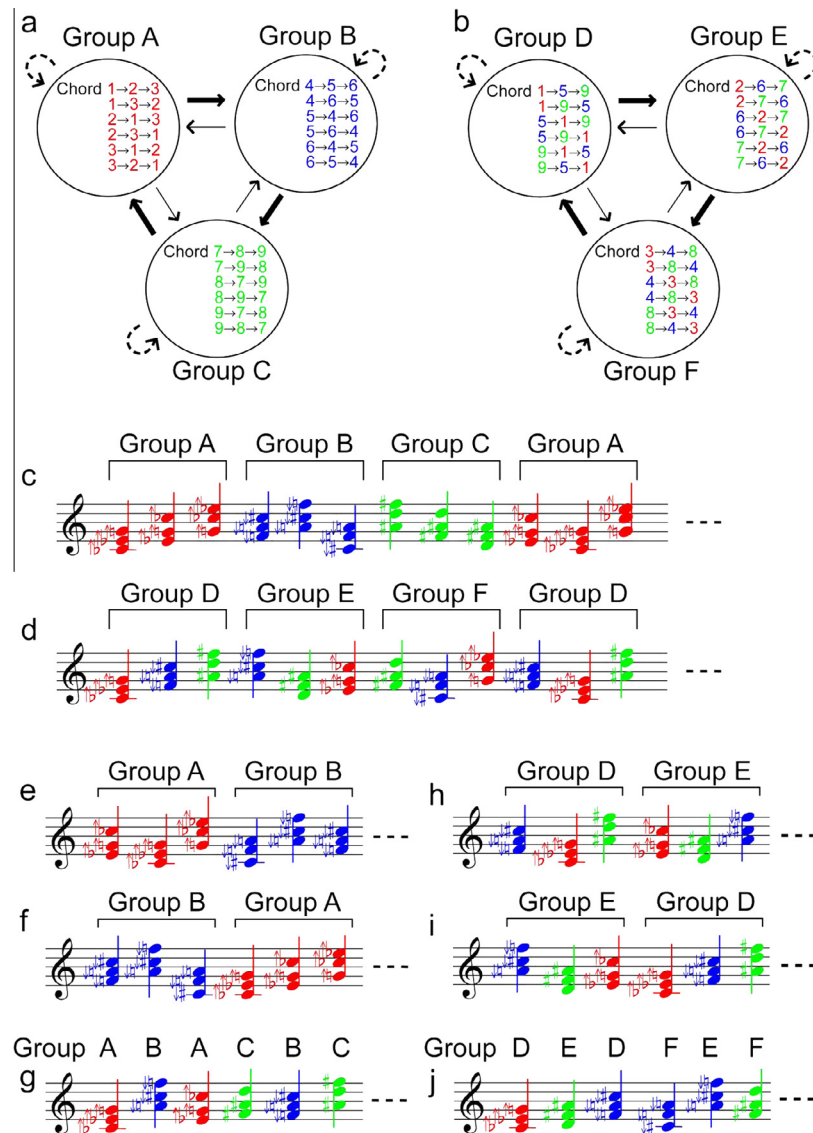


Fig. 2. Transition diagrams of the clustered pitch-classes sequences (a) and dispersed pitch-classes sequences (b). Each colour represents three distinct pitch classes. Each circle indicates triplet groups whose order was statistically defined such that the forthcoming group was statistically defined by the most recent group: the recursive transitions (dashed arrows) with 14% probability, and the clockwise (thick arrows) and counterclockwise transitions (thin arrows) with 80% and 6% probabilities, respectively. The 3-digit strings in the group circles indicate six possible chord triplets, which were randomly selected. Excerpts from the clustered (c) and dispersed (d) pitch-classes sequences during MEG measurements, and the Markov triplet (e), random triplet (f), and random chord series (g) in the behavioural tests of the clustered pitch-classes session and the Markov triplet (h), random triplet (i), and random chord series (j) in the behavioural tests of the dispersed pitch-classes session. They were written in musical scores with microtonal accidental notation, corresponding 292-Hz tone in the present study to the note D3. Microtonal accidentals of upward and downward arrows, which indicate a sixth of a whole tone higher and lower, respectively, are additionally applied to ordinary half tone accidentals: “ $\uparrow\sharp$ ” and “ $\downarrow\flat$ ” indicate the displacement of a scale degree by two sixths of a whole tone up or down, respectively; “ $\uparrow\sharp$ ” and “ $\downarrow\flat$ ” indicate the displacement of a scale degree by a sixth of a whole tone up or down, respectively.

order of the triplet groups was relaxed from Markovian to random. In other words, the random triplet series was a concatenation of chord triplets that belonged to either group of A, B or C in the clustered and D, E or F in the dispersed pitch-classes sessions, while the order of the triplet groups was not under Markovian regulation but randomised. In the third types of the clustered (random chord series; Fig. 2g) and dispersed (Fig. 2j) pitch-classes sessions, both of the between-triplet and within-triplet constraints were relaxed. That is to say, the nine chords were randomly ordered under the constraint that the same chord did not appear consecutively.

The behavioural tests following MEG measurement in each session were completed within six minutes per session per participant. The logit transformation was applied to normalize the ratios of answering that chord series sounded familiar (familiarity ratios). We conducted analysis of variance (ANOVA) with the logit

values of the familiarity ratios for each session. Bonferroni-corrected post-hoc tests were conducted for further analysis. Statistical significance levels were set at $p = 0.05$ for all analyses.

2.5. MEG data analysis

2.5.1. Artefact rejection

The procedure of MEG data analysis in the present study followed our previous studies (Daikoku et al., 2014, 2015). We used the original software implemented in the Neuromag system for the MEG data analysis. Epochs with artefacts that exceeded 3 pT/cm or 3 pT for any MEG channel were excluded from analyses. Contamination from environmental noise was reduced by using the temporally extended signal space separation method with a buffer length of 10 s and a correlation limit of 0.980 (Taulu & Hari, 2009).

2.5.2. Source analysis

For each participant, the responses to all the chords in both of the clustered and dispersed pitch-classes sequences were averaged to acquire reliable equivalent current dipoles (ECDs). The averaged responses were filtered offline with a 2–40 Hz band-pass. The cut-off frequency of 2 Hz for high-pass was selected to eliminate a residual fluctuation in the finite averaging data, considering that infinite averaging of continuous data at every SOA of 0.5 s eliminates signals below 2 Hz. The baseline for the magnetic signals in each MEG channel was defined by the mean amplitude in the

pre-stimulus period from –100 to 0 ms. The analysis window was defined as 0–500 ms. Using the 66 temporal channels for each participant (Fig. 3a), the ECDs were separately estimated with a peak magnitude during 40–80 ms for P1m, 80–140 ms for N1m, and 160–240 ms for P2m after sound onsets. Participants who demonstrated poor ECD estimation, with a goodness-of-fit below 75% in either the left or right hemisphere, were discarded from further analyses. Consequently, learning effects on the P1m, N1m, and P2m components were studied in 12, 10, and 12 participants, respectively.

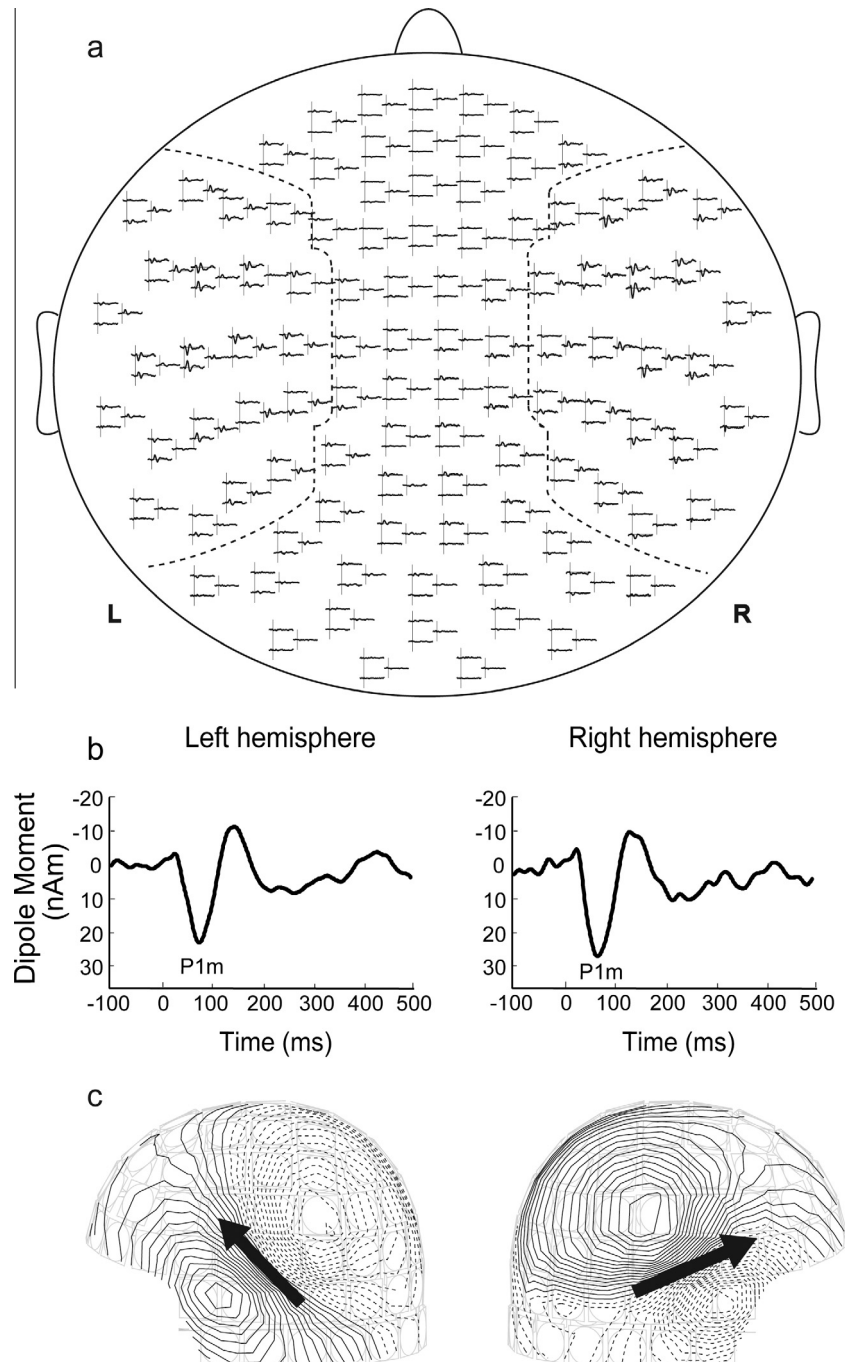


Fig. 3. A sensor layout of the averaged magnetic responses to all the chords presented during the measurement session from a representative participant. The selected 66 channels in each hemisphere for ECD estimation have dashed line borders (a). Source-strength waveforms for the P1m responses (b), which were calculated from the ECD for the averaged P1m responses as a template, and isofield contour maps for the P1m peak amplitude in each hemisphere (c). Outflux (solid lines) and influx (dashed lines) are stepped by 20 fT/cm. The black arrows represent the ECDs for P1m.

2.5.3. Source-strength waveforms

To confirm time course of learning effects from the neuromagnetic response series during the entire 1080-chord sequence, the responses to all of the initial, middle, and final chords in the triplets (refer to responses to the triplets, hereafter) in each 1080-chord sequence were averaged separately for the former and latter halves of each 1080-chord sequence, and for the clockwise and counterclockwise transitions with higher and lower transitional probabilities, respectively. The responses to the triplets transitioned from the same group were discarded from further analysis. In other words, in the former half, the responses to the triplets that appeared with the clockwise and counterclockwise transitions (i.e., higher and lower transitional probabilities, respectively) were separately averaged from the beginning to the midpoint of the sequence until the number of averaging reached the same count of 33 chords (11 times in each initial, middle, and final chord within a triplet). In the latter half, the same method of separate averaging was conducted from the midpoint to the end of the sequence. We analysed each response that was averaged 33 times as an event.

The source-strength waveforms for the P1m, N1m and P2m responses in each event were separately calculated using the ECDs for P1m, N1m and P2m, respectively (Fig. 3). In each sequence, we performed a 2 (half: former and latter) \times 2 (hemisphere: right and left) \times 2 (probability: high and low) repeated-measures ANOVA with the peak amplitude and latency of the source strength of the P1m, N1m and P2m. Bonferroni-corrected post-hoc tests were

conducted for further analysis. Statistical significance levels were set at $p = 0.05$ for all analyses.

3. Results

3.1. Behavioural data

For the clustered pitch-class session, two-tailed t -test results indicated that familiarity ratios were significantly above chance for the Markov triplet ($t[13] = 3.43, p = 0.004$) and random triplet series ($t[13] = 3.43, p = 0.005$) (Fig. 4). The ANOVA for the chord series detected a significant difference ($F[2,26] = 8.70, p = 0.001$). Bonferroni-corrected post-hoc tests revealed that the familiarity ratios were significantly higher for the Markov triplet ($p = 0.005$) and random triplet series ($p = 0.001$) compared to the random chord series. For the dispersed pitch-class session, two-tailed t -test results indicated that the familiarity ratios were not significant for any chord series. No additional significant results were detected in the behavioural tests.

3.2. MEG data

All of the participants raised their right hands at every silent period in the chord sequences and paid attention to the chord sequences. All of the peak amplitudes and latencies of the source-strength waveforms for the P1m, N1m, and P2m responses are shown in Table 2.

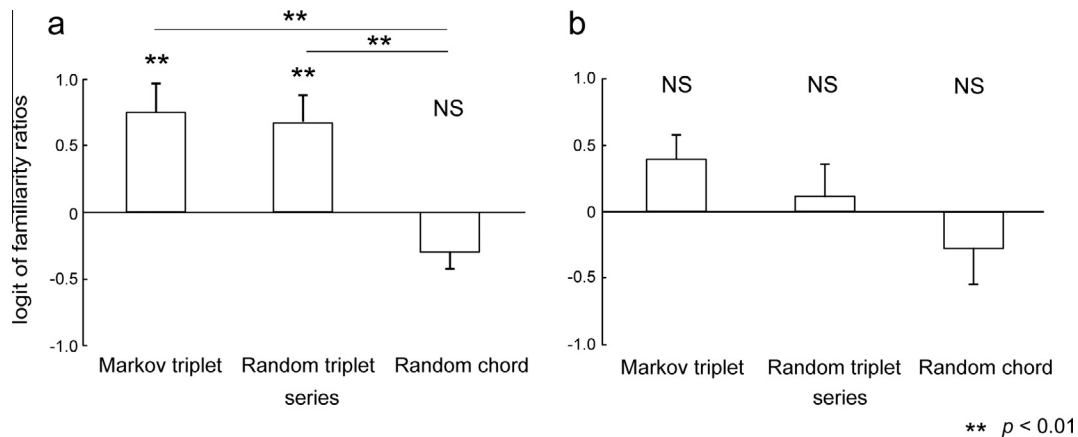


Fig. 4. The logit of the ratio of responses that indicated that the chord series sounded familiar for each chord series in the clustered pitch-classes sequences (a) and dispersed pitch-classes sequences (b). The familiarity ratio at logit = 0 corresponds to the chance level of 50%. The bars indicate the standard error of the mean.

Table 2

The peak amplitudes and latencies of the source strength for the P1m, N1m, and P2m responses.

	Sequence	Hemisphere	Clustered pitch-classes sequences				Dispersed pitch-classes sequences			
			Left		Right		Left		Right	
			Probability	High	Low	High	Low	High	Low	High
Former	P1m	Amplitude	24.7 \pm 5.0	18.6 \pm 2.3	13.8 \pm 3.6	13.5 \pm 3.1	22.9 \pm 3.8	23.0 \pm 5.0	10.1 \pm 2.1	13.9 \pm 3.4
		Latency	72.2 \pm 4.0	77.1 \pm 4.9	72.3 \pm 4.5	75.6 \pm 4.8	64.2 \pm 5.2	64.1 \pm 2.8	70.6 \pm 2.3	70.8 \pm 3.8
	N1m	Amplitude	17.7 \pm 4.1	14.6 \pm 3.5	21.0 \pm 36.0	17.1 \pm 4.5	11.8 \pm 5.0	9.2 \pm 3.0	22.2 \pm 5.5	16.2 \pm 3.0
		Latency	117.5 \pm 4.6	121.3 \pm 7.9	116.2 \pm 7.2	113.5 \pm 7.1	112.9 \pm 4.8	115.0 \pm 6.9	117.6 \pm 4.8	117.7 \pm 7.9
	P2m	Amplitude	19.2 \pm 4.0	21.8 \pm 5.9	14.7 \pm 4.5	17.2 \pm 5.3	16.1 \pm 3.6	18.3 \pm 4.2	16.7 \pm 5.5	14.8 \pm 2.2
		Latency	206.2 \pm 7.2	198.6 \pm 5.7	201.8 \pm 6.0	192.2 \pm 6.3	202.3 \pm 4.1	199.5 \pm 5.5	203.0 \pm 5.3	196.8 \pm 6.1
Latter	P1m	Amplitude	11.4 \pm 2.2	18.7 \pm 2.9	9.1 \pm 2.1	14.1 \pm 3.3	17.2 \pm 2.8	15.1 \pm 5.1	12.3 \pm 2.0	13.6 \pm 2.4
		Latency	61.3 \pm 8.0	74.8 \pm 6.8	71.6 \pm 5.0	78.8 \pm 3.7	74.8 \pm 3.9	67.6 \pm 5.9	76.5 \pm 3.0	71.5 \pm 5.2
	N1m	Amplitude	10.4 \pm 3.1	11.9 \pm 2.9	16.5 \pm 3.4	13.0 \pm 4.5	13.6 \pm 2.9	11.9 \pm 3.1	13.0 \pm 2.5	9.2 \pm 3.6
		Latency	117.4 \pm 6.8	109.3 \pm 10.2	120.5 \pm 5.7	123.6 \pm 7.2	118.1 \pm 9.7	113.8 \pm 8.4	116.4 \pm 9.8	114.7 \pm 8.1
	P2m	Amplitude	21.0 \pm 3.5	16.8 \pm 5.4	12.8 \pm 3.7	13.9 \pm 4.2	21.3 \pm 5.1	18.6 \pm 5.2	14.3 \pm 3.3	16.6 \pm 5.8
		Latency	200.4 \pm 7.0	192.8 \pm 5.2	200.2 \pm 6.2	202.5 \pm 5.4	200.1 \pm 5.2	194.4 \pm 4.9	201.2 \pm 5.9	200.3 \pm 5.3

Mean \pm standard error of the mean (amplitude: nAm, latency: ms).

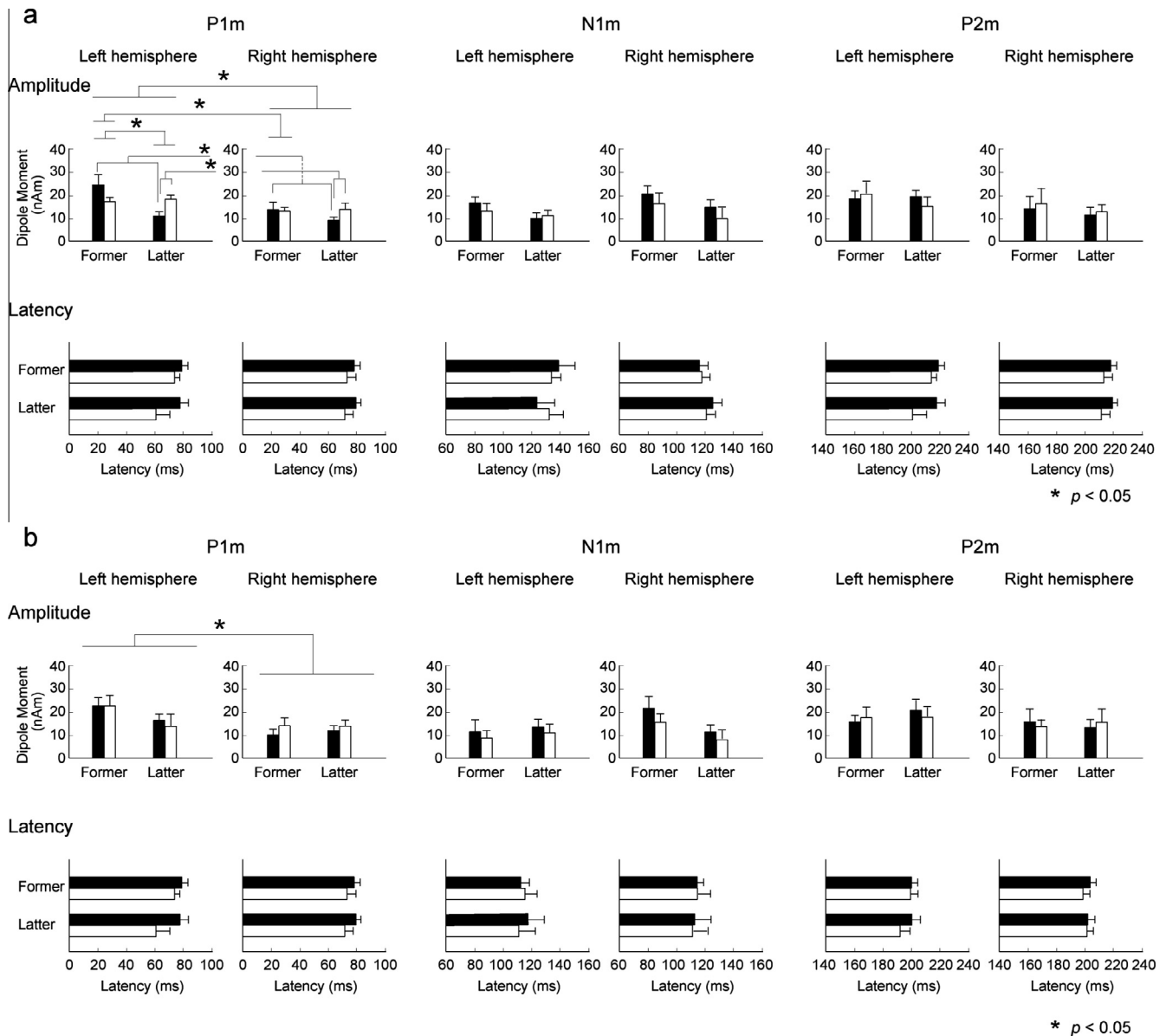


Fig. 5. The peak amplitudes and latencies of the source-strength waveforms for the P1m, N1m, and P2m responses in the clustered pitch-classes sequences (a) and dispersed pitch-classes sequences (b). The source-strength waveforms for the P1m, N1m, and P2m were separately calculated from the ECDs for P1m, N1m, and P2m, respectively. The closed bars represent responses to chords that had higher transitional probability, and the open bars represent responses to chords that had lower transitional probability. The error bars indicate the standard error of the mean. The asterisks indicate the main effects in the ANOVA and the simple main effects in the post-hoc tests.

3.2.1. Clustered pitch-classes sequences

The half-probability interaction of the P1m peak amplitude was significant ($F[1,11] = 4.99$, $p = 0.047$, partial $\eta^2 = 0.31$). The P1m responses to the triplets with higher transitional probability in the latter half were significantly decreased compared with those in the former half ($p = 0.022$), while no significance was detected in the P1m responses to the triplets with lower transitional probability ($p = 0.87$). In the latter half, the peak amplitudes were significantly lower in the P1m responses to the triplets with higher transitional probability than those with lower transitional probability ($p = 0.024$) (Fig. 6).

The main hemisphere effect on the P1m peak amplitudes were significant ($F[1,11] = 6.31$, $p = 0.029$, partial $\eta^2 = 0.37$). The P1m peak amplitudes were significantly higher in the left hemisphere than in the right hemisphere (Fig. 5).

The main half effect on the P1m peak amplitude was significant ($F[1,11] = 5.06$, $p = 0.046$, partial $\eta^2 = 0.32$). The P1m peak

amplitudes were significantly higher in the former half than in the latter half (Fig. 5).

The hemisphere-half interaction of the P1m peak amplitude was significant ($F[1,11] = 4.93$, $p = 0.048$, partial $\eta^2 = 0.31$). In the former half, the P1m peak amplitudes were significantly higher in the left hemisphere than in the right hemisphere ($p = 0.021$). In the left hemisphere, the P1m peak amplitudes in the latter half were significantly decreased compared with the former half ($p = 0.018$).

No other significant differences in the peak amplitudes and latencies of the source strength for the P1m, N1m, and P2m responses were detected.

3.2.2. Dispersed pitch-classes sequences

The main hemisphere effect on the P1m peak amplitudes were significant ($F[1,11] = 18.46$, $p = 0.0010$, partial $\eta^2 = 0.63$). The P1m peak amplitudes were significantly higher in the left hemisphere than in the right hemisphere (Fig. 5). There were no other

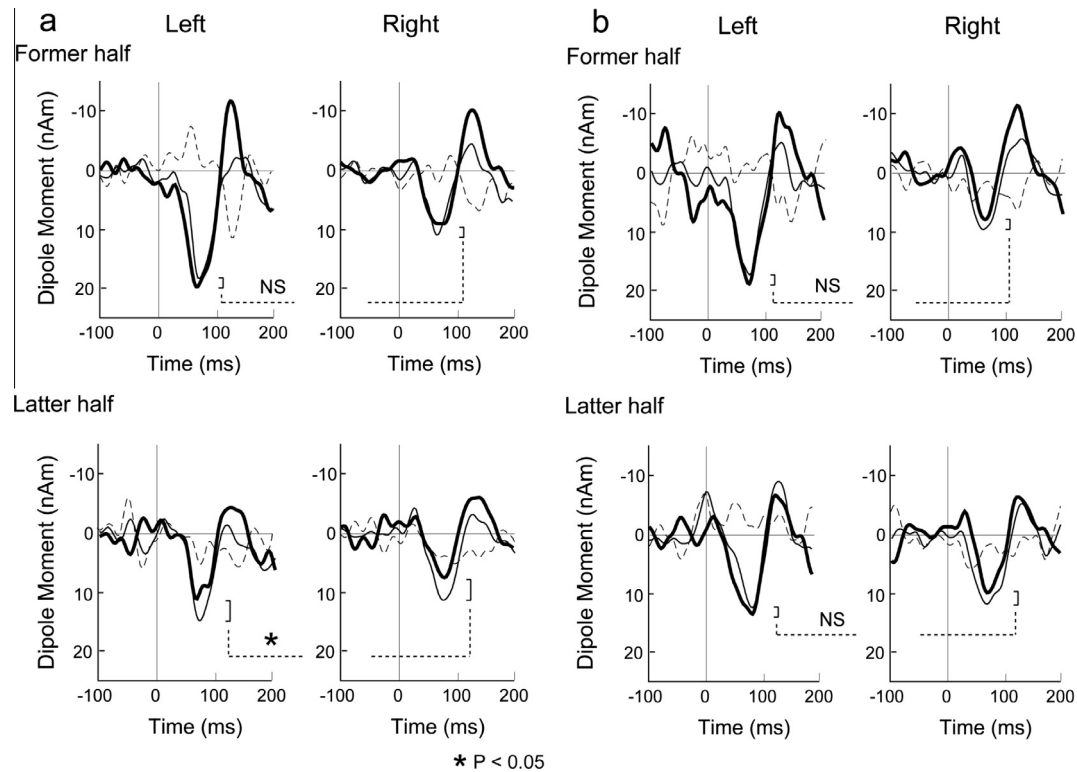


Fig. 6. Grand-averaged source-strength waveforms for the P1m responses to chords ($N = 12$) in the clustered pitch-classes sequences (a) and the dispersed pitch-classes sequences (b). The responses in the left and right hemispheres are located on the left and right sides, respectively. The solid lines represent the averaged responses to the triplet transitioned in clockwise direction with higher probability. The thin lines represent averaged responses to the triplet transitioned in counterclockwise direction with lower probability. The dashed lines represent difference waveforms.

significant differences in the peak amplitudes and latencies of the source strength for the P1m, N1m, and P2m responses.

4. Discussion

In previous studies on the statistical learning of single tone sequences, the statistical learning effect was detected as a significant reduction of the amplitude of the evoked responses to the tones that appeared with higher transitional probability compared to those with lower probability. The reduction was reflected not only in the N1 (N1m) response (Abla et al., 2008; Daikoku et al., 2014, 2015; Francois & Schön, 2010; Koelsch et al., 2016; Sanders, Ameral, & Sayles, 2009) but also in the P1m responses (Paraskevopoulos et al., 2012). Thus, the statistical learning of monophonic auditory sequences has been indexed by the evoked responses. Interestingly, in the present study, the statistical learning of the chord sequence was reflected in the P1m responses. The P1m responses, which exhibited no differences in the former half of the sequence, were significantly reduced relative to the triplets that appeared with higher transitional probability in the latter half of the clustered pitch-classes sequences, although such differences were not pronounced in the dispersed pitch-classes sequences. Additionally, in the clustered pitch-classes sequences, the P1m responses to the triplets with higher transitional probability in the latter half were significantly decreased compared with those in the former half, whereas this significance could not be detected in the P1m responses to the triplets with lower transitional probability. This interactive response reduction may correspond to the time course of learning achievement.

It may be important to note that the significant difference detected in the clustered pitch-classes sequences did not indicate the learning of triplet boundaries or rhythms, which had reasonably been assumed beforehand because such pitch-class changes at

every triplet should sound apparent to the listeners. In the clustered pitch-classes sequences, the nine chords in the same pitch classes were grouped into three (Table 1a), and the Markov transition occurred between the three groups (Fig. 2). The two possible transitions in clockwise and counterclockwise direction were compared in the present study. The significant response reduction detected in the clustered pitch-classes sequences indicates that the listeners learned the direction of atonal pitch-class transitions. Because the order of the three chords in triplets were randomised so that the listeners hardly anticipated the forthcoming triplets, the listeners may have recognized the chords in the same pitch class as an abstract-featured harmony and learned directional transition in the atonal chord sequence, regardless of chord orders or inversions within the same pitch-class triplets.

The behavioural results were consistent with the neurophysiological results. In the clustered pitch-class session, the familiarity ratios were significantly above chance for the Markov triplet series, in which chords were transitioned with the same statistical constraints as the chord sequence during MEG measurement. The significantly high familiarity ratios were concordant with the fact that the statistical learning of chord sequences was detected in neuro-magnetic responses. The familiarity ratios for the random triplet series, in which the order of the chord triplets was relaxed from Markov to pseudo-random, were also significantly above chance. The significant familiarity ratios support the former notion that the triplets in the same pitch class were recognized as one unit by the listeners. In the dispersed pitch-class session, the familiarity ratios were not significant for any type of chord series, corresponding to the neurophysiological results in the chord sequence with less regularity in pitch-class distribution. The neural processing underlying pitch-class perception, which has been interpreted as one of the domain-specific processing components in music, may

facilitate the domain-general statistical learning of chord sequences.

The present study did not detect a statistical learning effect on the N1m and P2m responses, although previous research detected an effect on the N1m and P2m as well as P1m (Daikoku et al., 2014, 2015; Furl et al., 2011; Koelsch et al., 2016; Paraskevopoulos et al., 2012). From the viewpoint of signal processing, infinite averaging of continuous data at every SOA of 0.5 s eliminates signals below 2 Hz in frequency. The relatively short SOA of 0.5 s corresponding to the applied high-pass filter of 2 Hz may be critical to refractory recovery of the later components such as P2m. However, we gave priority to the speed of chord progression around Allegretto – Allegro for emulating a habitual meter of contemporary music. This stimulus presentation speed may have relatively enhanced the earlier component of P1m. The maximum rising slope of the envelope of the stimulus waveforms, which was double that of the common 10-ms linear rise, may have synchronized neural activity and attributed to the build-up of the P1m response with shorter duration than the later components. The spectrum-rich auditory stimulus is another key to elicit the P1 component. The P1 (P1m) has been recorded using the paired-click paradigm to study sensory gating function (Adler et al., 1982). The transient clicks with short duration contain a broad frequency spectrum and recruit more neural activity with maximum synchrony than single sine tones. The complex tones composed of three sine tones with equal intensity used in the present study may have helped elicit the P1m responses, otherwise showing relatively high inter-individual variability among other evoked components.

According to prior research, learning effects on the P1 responses were not correlated with learning effects on the other components of neural responses (Boutros & Belger, 1999; Boutros et al., 1995; Kisley, Noecker, & Guinther, 2004). For example, previous studies suggested that statistical learning effects on P1 responses were more detectable in musicians than non-musicians, whereas those on N1 responses did not differ between musicians and non-musicians (Paraskevopoulos et al., 2012). This implies that the effects on P1 responses may depend on music expertise and specialized training experience (Kizkin, Karlidag, Ozcan, & Ozisik, 2006; Wang, Staffaroni, Reid, Steinschneider, & Sussman, 2009). Another implication is that the P1 responses are by nature more relevant to musical sequences than other components, especially for the initial learning phase. In the present study, although participants were not grouped based on music expertise, we equalized learning achievement across participants by using atonal chord sequences of augmented triads in a nine-tone equal temperament, which have never been adopted in tonal music. Furthermore, we devised original rules for chord sequences; thus, it should be unlikely that participants had prior experience. Despite having little knowledge of the chord sequences used in the present study, learning effects were solely reflected in the P1m responses during statistical learning of chord sequences. The pitch-class perception could be reflected in the P1 responses, and may facilitate the statistical learning of chord sequences. In previous research, in which statistical learning effects were detected in the N1m and P2m but not P1m, stimuli were single tone sequences and therefore unlikely to be exposed to the effects of pitch-class perception (Abla et al., 2008; Daikoku et al., 2014, 2015; Furl et al., 2011).

The P1 and N1 responses to auditory stimuli are considered to be generated in the auditory cortex with different topographies (Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001). The P1 is generated in the lateral part of the primary auditory cortex, and the N1 is generated in the secondary auditory cortex, which is more lateral and anterior to the primary auditory cortex, and in the planum temporale. A previous study reported that the earlier auditory cortical responses, which peaked at approximately

20–80 ms, may be attributed to parallel thalamo-cortical connections, or to cortico-cortical connections between the primary auditory cortex and the superior temporal gyrus (Yvert et al., 2001). Further studies are needed to clarify the specific attributes of the P1 responses in statistical learning.

In both sessions, P1m amplitudes were significantly higher in the left than the right hemispheres and the amplitudes in the left hemisphere were significantly decreased in the latter half. In contrast, we did not detect significant laterality in the N1m. According to previous studies, the P1 response could be attenuated with a repetition of an identical stimulus due to the adaptation of auditory cortical neurons (Erwin & Buchwald, 1986; Javitt, 2000; Lu, Williamson, & Kaufman, 1992). However, if auditory responses to stimulus repetitions with changing frequencies had been investigated, P1 responses could be increased with repetition (Dyson, Alain, & He, 2005; Haenschel, Vernon, Dwivedi, Gruzeliier, & Baldeweg, 2005), particularly in the left hemisphere (Chakalov, Draganova, Wollbrink, Preissl, & Pantev, 2012). The difference in the behaviour of the P1 and N1 responses to stimulus repetition with changing frequencies suggests that the neural mechanisms underlying auditory adaptation that are reflected in the P1 and N1 may be partially different. Further studies are needed to examine the specific attributes of P1 responses in statistical learning and pitch-class perception.

The previous studies reported that the adaptation effects on the N1m and P2m responses to repeated stimuli were detected in different time scales and polarity (Ross & Tremblay, 2009). The participants attended two experimental sessions on different days. The N1m amplitudes were attenuated during each session and recovered between the two sessions. In contrast, the P2m amplitudes were fairly constant within a session, but increased from the first to the second session (Ross & Tremblay, 2009).

On the other hand, the participants who demonstrated poor ECD estimation, with a goodness-of-fit below 75% in either the left or right hemisphere, were discarded. Because there were uneven sample sizes for each component (P1m, N1m, and P2m components were studied in 12, 10, and 12 participants, respectively), we could not include the component of P1m, N1m, and P2m as a factor in an ANOVA. The results of the ANOVA in the present study was not sufficient to make a conclusion that the P1m can specifically reflect statistical learning of atonal chord sequences. In addition, we could not exclude the possibility that the fewer sample sizes of N1m might cause the insignificant results. Nevertheless, our results that significant difference in P1m amplitudes between high and low probabilities were detected in the latter half but not in the former half indicated that time course of statistical learning effects could be reflected in the neural responses.

The results of the present study are consistent with the general hypothesis that musical structure is acquired by both the domain-general and domain-specific learning mechanisms (Jonaitis & Saffran, 2009; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Tillmann et al., 2000). Sequential statistics, including transitional probabilities, may be cues that are useful to learning harmonic structure, which implies that domain-general statistical learning is an essential strategy for understanding music. Because sequential statistics, including transitional probabilities, can be useful cues to learning harmonic structure, domain-general statistical learning may be an essential strategy for understanding music. In contrast, because learning chord sequences that are organized by pitch class can be reflected in both neural and behavioural responses, domain-specific perceptions that underlie pitch class may facilitate chord sequence learning in music.

In conclusion, our study suggests that domain-specific perception underlying the pitch class may facilitate domain-general chord sequence statistical learning. Furthermore, P1m responses may be a useful probe for the statistical learning of chord

sequences and the perception underlying pitch class. To the best of our knowledge, this is the first study to demonstrate that the statistical learning of novel syntactic structures of atonal chord sequences can be reflected in both behavioural and neural responses. Understanding how learners acquire music-syntactic structures may help illuminate the nature and roles of domain-general and domain-specific learning mechanisms.

References

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, 20, 952–964.
- Adler, L. E., Pachtman, E., Franks, R. D., Pecevich, M., Waldo, M. C., & Freedman, R. (1982). Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia. *Biological Psychiatry*, 17, 639–654.
- Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, 125, 238–251.
- Boutros, N. N., & Belger, A. (1999). Midlatency evoked potentials attenuation and augmentation reflect different aspects of sensory gating. *Biological Psychiatry*, 45, 917–922.
- Boutros, N. N., Torello, M. W., Barker, B. A., Tueting, P. A., Wu, S. C., & Nasrallah, H. A. (1995). The P50 evoked potential component and mismatch detection in normal volunteers: Implications for the study of sensory gating. *Psychiatry Research*, 57, 83–88.
- Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology: General*, 113, 394–412.
- Chakalov, I., Draganova, R., Wollbrink, A., Preissl, H., & Pantev, C. (2012). Modulations of neural activity in auditory streaming caused by spectral and temporal alternation in subsequent stimuli: A magnetoencephalographic study. *BMC Neuroscience*, 13, 1471–2202.
- Daikoku, T., Yatomi, Y., & Yumoto, M. (2014). Implicit and explicit statistical learning of tone sequences across spectral shifts. *Neuropsychologia*, 63, 194–204.
- Daikoku, T., Yatomi, Y., & Yumoto, M. (2015). Statistical learning of music- and language-like sequences and tolerance for spectral shifts. *Neurobiology of Learning and Memory*, 118, 8–19.
- Drobisch, M. W. (1855). Über musikalische Tonbestimmung und Temperatur. *Abhandlungen der Mathematisch-Physischen Klasse der Sächsischen Akademie der Wissenschaften* 2.
- Dyson, B. J., Alain, C., & He, Y. (2005). I've heard it all before: Perceptual invariance represented by early cortical auditory-evoked responses. *Cognitive Brain Research*, 23, 457–460.
- Erwin, R. J., & Buchwald, J. S. (1986). Midlatency auditory evoked responses: Differential recovery cycle characteristics. *Electroencephalography and Clinical Neurophysiology*, 64, 417–423.
- Francois, C., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, 21, 928–932.
- Francois, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, 308, 122–128.
- Frensch, P. A., & Buchner, A. (1999). Domain-general versus domain-specificity in cognition. *The Nature of Cognition*, 137–172.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, 50, 259–281.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, 1, 183–192.
- Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., & Griffiths, T. D. (2011). Neural prediction of higher-order auditory sequence statistics. *Neuroimage*, 54, 2267–2277.
- Guo, S., & Koelsch, S. (2015). The effects of supervised learning on event-related potential correlates of music-syntactic processing. *Brain Research*. <http://dx.doi.org/10.1016/j.brainres.2015.01.046>.
- Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience*, 25, 10494–10501.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Jackendoff, R., & Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100, 33–72.
- James, C. E., Cereghetti, D. M., Roulet, T. E., & Oechslin, M. S. (2015). Electrophysiological evidence for a specific neural correlate of musical violation expectation in primary-school children. *Neuroimage*, 5(5). <http://dx.doi.org/10.1016/j.neuroimage.2014.09.047> (Epub 2014 Sep 30).
- Javitt, D. C. (2000). Intracortical mechanisms of mismatch negativity dysfunction in schizophrenia. *Audiology and Neuro-Otology*, 5, 207–215.
- Jonaitis, E. M., & Saffran, J. R. (2009). Learning harmony: The role of serial statistics. *Cognitive Science*, 33, 951–968.
- Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, 159–201.
- Keiler, A. (1978). Berstein's the unanswered question and the problem of musical competence. *The Musical Quarterly*, 64, 195–222.
- Kim, S. G., Kim, J. S., & Chung, C. K. (2011). The effect of conditional probability of chord progression on brain response: An MEG study. *PLoS One*, 6. <http://dx.doi.org/10.1371/journal.pone.0017337>.
- Kisley, M. A., Noecker, T. L., & Gunther, P. M. (2004). Comparison of sensory gating to mismatch negativity and self-reported perceptual phenomena in healthy adults. *Psychophysiology*, 41, 604–612.
- Kizkin, S., Karlıdag, R., Ozcan, C., & Ozisik, H. I. (2006). Reduced P50 auditory sensory gating response in professional musicians. *Brain and Cognition*, 61, 249–254.
- Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, 46, 179–190.
- Koelsch, S., Busch, T., Jentschke, S., & Rohrmeier, M. (2016). Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Scientific Reports*, 6, 19741.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schroger, E. (2000). Brain indices of music processing: "Non-musicians" are musical. *Journal of Cognitive Neuroscience*, 12, 520–541.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203–205.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: Evaluation and topography of the middle latency components. *Electroencephalography and Clinical Neurophysiology*, 92, 204–214.
- Lu, Z. L., Williamson, S. J., & Kaufman, L. (1992). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain Research*, 572, 236–241.
- Markov, A. A. (1971). Extension of the limit theorems of probability theory to a sum of variables connected in a chain. *Markov chains* (Vol. 1). John Wiley and Sons (Reprinted in Appendix B of: R. Howard D. Dynamic Probabilistic Systems).
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, 50, 341–349.
- Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10, 717–733.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6, 688–691.
- Poon, H., & Domingos, P. (2007). Joint inference in information extraction. In *Proceedings of the twenty-second national conference on artificial intelligence* (pp. 913–918). Vancouver, Canada: AAAI Press.
- Poon, H., & Domingos, P. (2008). Joint unsupervised coreference resolution with Markov Logic. In *Proceedings of the 2008 conference on empirical methods in natural language processing* (pp. 650–659). Honolulu, Hawaii: Association for Computational Linguistics.
- Richardson, M., & Domingos, P. (2006). Markov logic networks. *Machine Learning*, 62 (1–2), 107–136.
- Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, 248, 48–59.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70, 27–52.
- Saffran, J. R., Loman, M. M., & Robertson, R. R. (2001). Infant long-term memory for music. *Annals of the New York Academy of Sciences*, 930, 397–400.
- Sanders, L. D., Ameral, V., & Sayles, K. (2009). Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia*, 47, 1183–1186.
- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, 106, 975–983.
- Schön, D., & François, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontier in Psychology*, 2, 167.
- Singla, P., & Domingos, P. (2006). Entity resolution with markov logic. In *ICDM '06: Proceedings of the sixth international conference on data mining* (pp. 572–582). Washington, DC, USA: IEEE Computer Society.
- Smith, J. D., & Melara, R. J. (1990). Aesthetic preference and syntactic prototypicality in music: 'Tis the gift to be simple. *Cognition*, 34, 279–298.
- Taulu, S., & Hari, R. (2009). Removal of magnetoencephalographic artifacts with temporal signal-space separation: Demonstration with single-trial auditory-evoked responses. *Human Brain Mapping*, 30, 1524–1534.
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107, 885–913.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100, 73–99.
- Wang, W., Staffaroni, L., Reid, E., Steinschneider, M., & Sussman, E. (2009). Effects of musical training on sound pattern processing in high-school students. *International Journal of Pediatric Otorhinolaryngology*, 73, 751–755.
- Whittall, A. (2008). *The Cambridge introduction to serialism*. New York: Cambridge University Press.
- Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cerebral Cortex*, 11, 411–423.