Concurrent Statistical Learning of Ignored and Attended Sound Sequences: An MEG Study

*Tatsuya Daikoku [1,2] * and Masato Yumoto [1]*

[1]Department of Clinical Laboratory, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, [2]Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

In an auditory environment, humans are frequently exposed to overlapping sound sequences such as those made by human voices and musical instruments, and we can acquire information embedded in these sequences *via* attentional and nonattentional accesses. Whether the knowledge acquired by attentional accesses interacts with that acquired by nonattentional accesses is unknown, however. The present study examined how the statistical learning (SL) of two overlapping sound sequences is reflected in neurophysiological and behavioral responses, and how the learning effects are modulated by attention to each sequence. SL in this experimental paradigm was reflected in a neuromagnetic response predominantly in the right hemisphere, and the learning effects were not retained when attention to the tone streams was switched during the learning session. These results suggest that attentional and nonattentional learning scarcely interact with each other and that there may be a specific system for nonattentional learning, which is independent of attentional learning.

Keywords: statistical learning, attention, auditory, Markov model, magnetoencephalography

INTRODUCTION

Statistical learning (SL) is a domain-general and automatic process that is innate to humans (Saffran et al., 1996; Perruchet and Pacton, 2006). By this process, the brain computes transitional probabilities (TPs) of sequential phenomena such as music and language without intention or awareness (Cleeremans et al., 1998), and incessantly updates acquired statistical knowledge to adapt to variable phenomena in environments (Daikoku et al., 2017c; Daikoku, 2018a,c,d).

Such SL effects have been observed in neurophysiological responses. For instance, the event-related potentials (ERPs) and magnetic fields (ERFs) represent a more sensitive method than behavioral responses (Schön and François, 2011; Paraskevopoulos et al., 2012; Koelsch et al., 2016). In a framework of predictive coding (Friston, 2005), when the brain codes TP distributions of a stimulus sequence, it expects a probable future stimulus with a high TP and inhibits the neural response to predictable external stimuli. Finally, the SL effects manifest as a difference in amplitudes between stimuli with lower and higher TPs. A body of studies detected SL effects on ERP/ERF such as P50 (Paraskevopoulos et al., 2012; Daikoku et al., 2016, 2017c; Daikoku and Yumoto, 2017), N100 (Sanders et al., 2002; Furl et al., 2011; Daikoku et al., 2014, 2015, 2017c), mismatch negativity (MMN; Koelsch et al., 2016; François et al., 2017; Moldwin et al., 2017), P200 (Cunillera et al., 2006; De Diego Balaguer et al., 2007; François and Schön, 2011; Furl et al., 2011), P300 (Batterink et al., 2015), and N400 components (Sanders et al., 2002; Cunillera et al., 2006, 2009; François and Schön, 2011; François et al., 2013, 2014). Compared

with later auditory responses, the earlier auditory responses that peak at 20–80 ms (e.g., P50) have been attributed to parallel cortico-cortical or thalamo-cortical connections between the primary auditory cortex and the superior temporal gyrus (Adler et al., 1982). Thus, suppression of an early component of auditory responses to stimuli with a higher TP in lower cortical areas can be regarded as a transient expression of prediction error that is suppressed by predictions from higher cortical areas in a top-down connection (Skoe et al., 2015).

Most neurophysiological studies on SL have investigated the SL of single-tone sequences. In real-world auditory environments, however, humans are simultaneously exposed to overlapping sound sequences such as those made by musical instruments and human voices. Even when we selectively attend to the important information and ignore the unimportant information in overlapping sounds, humans generally acquire the information through both attentional and nonattentional processes (Jimenez and Castor, 1999; Aizenstein et al., 2004; Daikoku and Yumoto, 2017; Yumoto and Daikoku, 2018). However, few neurophysiological studies have examined attentional and nonattentional SL when learners are simultaneously exposed to multiple streams of sequences. To understand the mechanisms underlying SL, which is considered to occur automatically regardless of attention (Perruchet and Pacton, 2006), it is important to investigate how concurrent SL of attended and ignored sequences is reflected in neural responses.

In studies addressing consciousness during learning, the learning system has been divided into implicit learning, which may be accomplished through unconscious and nonattentional learning processes, and explicit learning, which may be accomplished through conscious and attentional learning processes (Reber, 1989; Ellis, 2005, 2009; Daikoku and Yumoto, 2017). The earlier studies suggested that explicit and implicit knowledge could be acquired by different learning processes and that explicit knowledge cannot be transformed into implicit knowledge through practice (Hulstijn, 2002). In contrast, other researchers have demonstrated that implicit and explicit knowledge can interact with each other (DeKeyser, 2003, 2007; De Jong, 2005; Ellis, 2005, 2009). Thus, interactive mechanisms between implicit and explicit learning remain a matter for debate (Krashen, 1982; Hulstijn, 2002; Daikoku et al., 2017b, 2018; Daikoku, 2018b).

To understand the neural mechanisms underlying concurrent attentional and nonattentional SL of auditory sequences, the present study used magnetoencephalography (MEG), a modality that can clearly resolve signals produced by the auditory cortices located bilaterally in the temporal lobes. We investigated how concurrent SL of simultaneous sequences of auditory stimuli is reflected in neuromagnetic responses and how the two forms of SL neurophysiologically interact with each other. MEG was recorded while participants listened to a dyad sequence (two-note chord). The dyad sequence can also be regarded as two types of auditory sequences consisting of low- and high-voice sequences in a distinct Markov-chain relationship. During the last third of each sequence, however, the Markov chains controlling the low and high voices were exchanged. The subjects were instructed to ignore one of the two types of sequences but

to attend to the other. Given neural representations of SL effects, we hypothesized that if subjects could concurrently perform SL of the two sequences, a dyad that consisted of two frequent tones with higher TP should lead to the lowest response amplitudes, while a dyad that consisted of two rare tones with lower TP should lead to the highest response amplitudes. Furthermore, if the statistical knowledge of attended sequences and that of ignored sequences cannot be transformed from one type to the other, the SL effect should disappear when the sequential regulation of the high and low voices is exchanged in the final third of the sequence. In contrast, if statistical knowledge of the attended and ignored sequences can interact and be transformed, the SL effect should remain even when the sequential regulation of the low and high voices is exchanged.

MATERIALS AND METHODS

Participants

Fifteen right-handed (57.9–100 in Edinburgh handedness, Oldfield, 1971) subjects without neurological and audiological disabilities participated (age range: 24–36 years, seven females, no absolute pitch). The present study was approved by the Ethics Committee of The University of Tokyo. All subjects were informed about this experiment including protection and safety of personal data, then provided written informed consent. The present study was conducted based on the guidelines and regulations.

Stimuli

We used the same stimuli as those in our previous study (Daikoku and Yumoto, 2017). The eight complex tones consisted of four high and low pitches each based on a five-tone equal temperament ($100 \times 2^{(n-1)/5}$ Hz, high: $n = 11$–14: 400, 459, 528, and 606 Hz; low: $n = 1$–4: 100, 115, 132, and 152 Hz; duration 350 ms with rise/fall of 10/150 ms; 80 dBSPL intensity and binaural presentation). The sequence consisted of 1,092 repetitions of two-tone chords (SOA = 500 ms), each of which consisted of a high and low pitches within which the intervals were separated by more than one octave (**Figure 1**).

The order in which the high and low pitches was defined separately based on a second-order Markov model (Markov, 1971, reprinted) with the constraint that the probability of a forthcoming tone was statistically defined (80% for a tone; 6.67% for the other three tones) by the last two successive tones
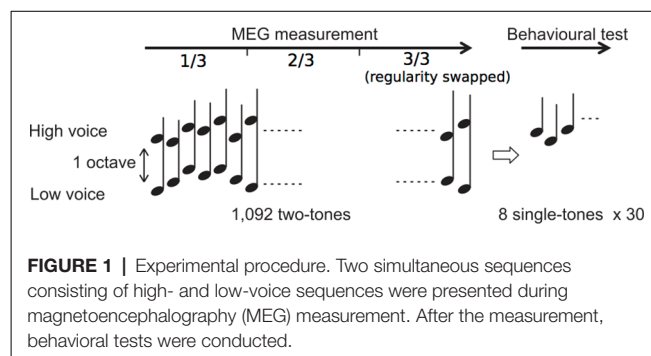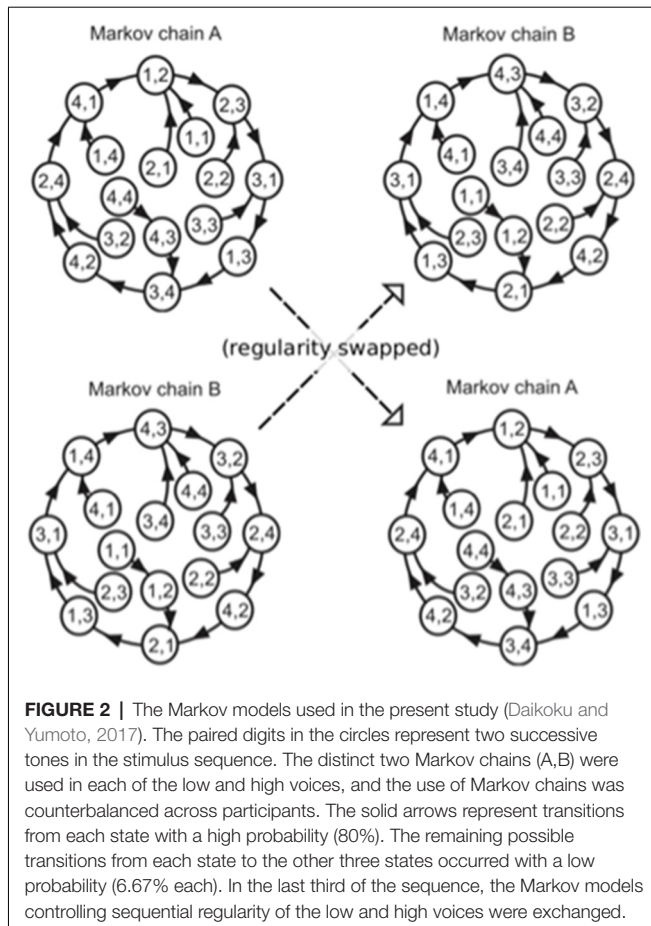

FIGURE 1 | Experimental procedure. Two simultaneous sequences consisting of high- and low-voice sequences were presented during magnetoencephalography (MEG) measurement. After the measurement, behavioral tests were conducted.

FIGURE 2 | The Markov models used in the present study (Daikoku and Yumoto, 2017). The paired digits in the circles represent two successive tones in the stimulus sequence. The distinct two Markov chains (A,B) were used in each of the low and high voices, and the use of Markov chains was counterbalanced across participants. The solid arrows represent transitions from each state with a high probability (80%). The remaining possible transitions from each state to the other three states occurred with a low probability (6.67% each). In the last third of the sequence, the Markov models controlling sequential regularity of the low and high voices were exchanged.

(**Figure 2**). In the last third of the sequence, however, the Markov models controlling sequential regularity of the low and high voices were exchanged (**Figure 1**). The regularities of the Markov models were counterbalanced across subjects.

Experimental Protocol

Subjects listened to a 1,092-dyad sequence with MEG measurement and took a behavioral test immediately afterward. They were instructed to ignore one sequence but attend to the other. The assignment of the attended and ignored sequences was counterbalanced across the subjects. To distinguish between attended and ignored conditions, a silent period of 500 ms was pseudo-randomly inserted within every set of 40 successive tones in attended sequence. Before the session, the subjects were instructed to raise their right hands at every silent period in attended sequence. Thus by observing that all subjects correctly raised their right hands at every silent period, we were able to confirm that they continually paid attention to only attended sequence.

After the measurement, subjects were presented with 30 series each consisting of eight single tones. Subjects answered whether each eight-tone series sounded familiar or not. The 30 series of eight tones could be classified into three types, and the presentation order was randomized. In 10 series, tone stimuli were sequenced using the Markov model that was applied in

the last third of the ignored sequence (tone series A). In an additional 10 series, tone stimuli were sequenced according to the same Markov model as an attended sequence in last third of the sequence (tone series B). In the remaining 10 series, tones were pseudo-randomly ordered (random tone series). The behavioral test was completed within 6 min for each subject.

Measurement and Data Analysis

Measurement and analysis were conducted as in our previous studies (Daikoku and Yumoto, 2017). Selective response averaging was performed separately for the first, middle, and last thirds of the sequence. Responses to each chord were selectively averaged from the beginning of each first, middle, and last thirds of the sequence. They were also selectively averaged in each dyad stimulus: chord that consisted of two high-TP (i.e., frequent) tones in both attended and ignored sequences, chord that consisted of two low-TP (i.e., rare) tones in both attended and ignored sequences, chord that consisted of a frequent tone in attended sequence and a rare tone in ignored sequence, and chord that consisted of a rare tone in the attended sequence and a frequent tone in the ignored sequence. The averaged responses were filtered offline with a 2–40 Hz band-pass. The baseline for the magnetic signals in each MEG channel was defined by the mean amplitude in the pre-dyad period from -100 to 0 ms. The analysis window was defined as 0–500 ms. In addition to selective averaging, all responses (1,092-dyad stimuli) to the dyads were averaged in each subject, enabling us to evaluate reliability for individual components. Using the averaged responses to all 1,092-dyad stimuli, the P1m, N1m and P2m were separately modeled as single equivalent current dipoles (ECDs) in each hemisphere (Daikoku et al., 2017a). The ECDs were calculated from the averaged responses to all 1,092-dyad stimuli with a goodness of fit above 80% using the 66 temporal channels (44 gradiometers and 22 magnetometers) for each participant. The selected channel areas correspond to our previous studies (Daikoku et al., 2014, 2015, 2016, 2017a). Subjects who demonstrated poor ECD estimation, with a goodness-of-fit below 80% 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 13, 10, and 11 subjects, respectively. Because a lot of the goodness of fit in the ECDs for the N1m and P2m were less than 80%, they were excluded from the analyses in this study.

Using the ECDs, the source-strength for P1 m in each hemisphere were calculated based on selective response averaging. Then, we performed a 3 (portion: first, middle, and last) $\times$ 2 (hemisphere: right and left) $\times$ 4 (dyad stimulus: chord that consisted of two frequent tones in both attended and ignored sequences, chord that consisted of two rare tones in both attended and ignored sequences, chord that consisted of a frequent tone in attended sequence and a rare tone in ignored sequence, and chord that consisted of a rare tone in the attended sequence and a frequent tone in the ignored sequence) repeated-measures analysis of variance (ANOVA) with peak amplitude and the latency of the source-strength of P1m. Bonferroni-corrected *post hoc* tests were conducted for further analysis. Furthermore, we performed ANOVA with the logit
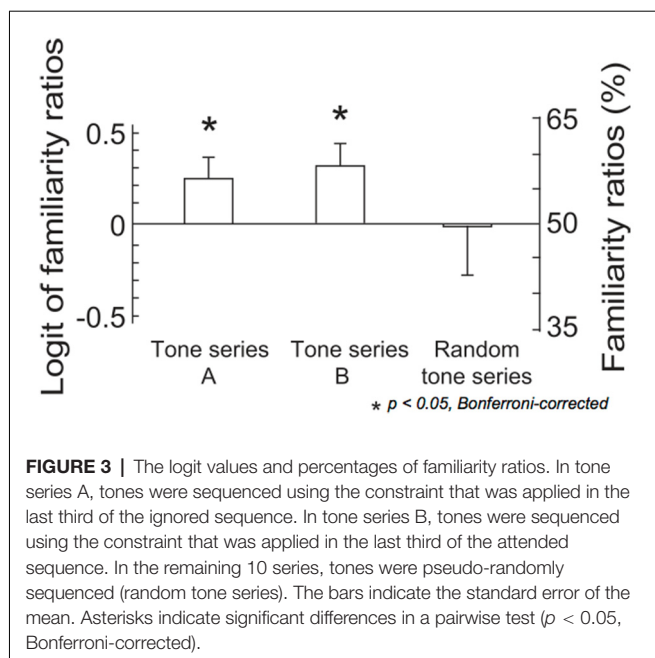
FIGURE 3 | The logit values and percentages of familiarity ratios. In tone series A, tones were sequenced using the constraint that was applied in the last third of the ignored sequence. In tone series B, tones were sequenced using the constraint that was applied in the last third of the attended sequence. In the remaining 10 series, tones were pseudo-randomly sequenced (random tone series). The bars indicate the standard error of the mean. Asterisks indicate significant differences in a pairwise test ($p < 0.05$, Bonferroni-corrected).

values of the familiarity ratios in behavioral test. Significance levels were set at $p = 0.05$ for all analyses. For further analysis, *post hoc* tests with Bonferroni correction were performed.

RESULTS

Behavioral Results

The results of two-tailed t-tests indicated that the familiarity ratios were significantly above chance level in both tone series A and tone series B (tone series A: $t_{(14)} = 2.30$, $p = 0.037$, tone series B: $t_{(14)} = 2.46$, $p = 0.028$; **Figure 3**). The ANOVA detected no significant results.

MEG Results

The averaged peak amplitudes and latencies of P1m responses are shown in **Figure 4**. The ANOVA detected that the main portion effect on the amplitudes was significant ($F_{(2,24)} = 3.74$, $p = 0.039$). The amplitudes in the last portion were significantly greater than those in the first portion ($p = 0.049$). The hemisphere-stimulus-portion interaction of the amplitudes was significant ($F_{(6,72)} = 2.32$, $p = 0.042$). In the middle and last portions, the amplitudes for the dyads that consisted of two frequent tones were significantly higher in the left than in the right hemispheres (middle: $p = 0.039$, last: $p = 0.036$). In the right hemisphere, the amplitudes for the dyads that consisted of two rare tones were significantly higher than those for the dyads that consisted of two frequent tones in the middle portion ($p = 0.028$). The results were consistent with a body of previous studies on SL: the brain learned TPs of the sequences, predicted a stimulus with a high TP (i.e., frequent stimuli), and inhibited the neural response to the stimuli with a high TP. The SL effects finally represent as a difference amplitudes between the stimuli with high and low TPs (François and Schön, 2011; François et al., 2013, 2017; Paraskevopoulos et al., 2012; Daikoku et al., 2014,
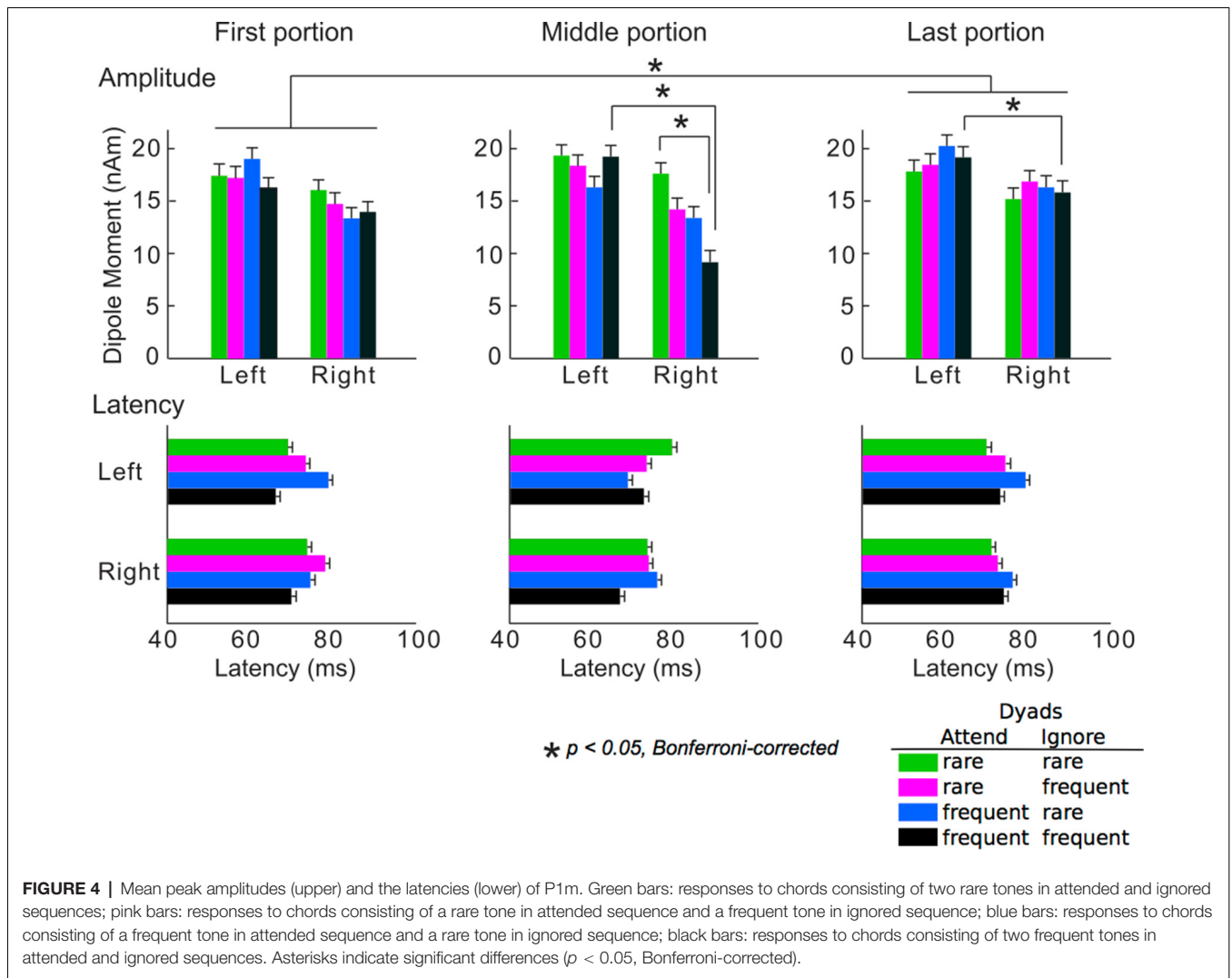
2015, 2016; Koelsch et al., 2016). These SL effects (i.e., difference amplitudes between the stimuli with high and low TPs), however, could not be detected after the Markov chains of the two sequences were exchanged in the last portion. This may suggest that SL effects cannot be retained when sequential regulations are exchanged. There was no significance in latency. No other significant differences were detected.

DISCUSSION

When the brain encodes the TP distributions of a stimulus sequence, humans expect a probable future stimulus with a high TP and inhibit the neural response to predictable external stimuli. In the end, the effects of SL manifest as a difference in amplitudes between stimuli with lower and higher TPs (Yumoto and Daikoku, 2016; Daikoku, 2018b). In the present study, subjects listened to two simultaneous sequences composed of tones with lower and higher TPs (i.e., rare and frequent tones, respectively). The subjects were instructed to ignore one of the two simultaneous sequences and to attend to the other. Based on the combinations of rare and frequent tones in the two simultaneous sequences, there were four types of dyads: dyads consisting of frequent tones in both sequences, dyads consisting of rare tones in both sequences, dyads consisting of a frequent tone in the attended sequence and a rare tone in the ignored sequence, and vice versa. If subjects were able to perform the SL of two sequences, and simultaneously predict stimuli with high TPs in both sequences, dyads consisting of two frequent tones should generate the lowest-amplitude responses, while those consisting of two rare tones should generate the highest-amplitude responses.

We found that, in the right hemisphere, neural responses to dyads consisting of two rare tones in ignored and attended sequences were significantly greater than those to dyads consisting of two frequent tones in ignored and attended sequences. These results suggested that the subjects were able to learn the statistics of the two sequences simultaneously and that SL of a sequence of two-tone dyads may be right-hemisphere dependent. This result is in agreement with previous studies that have reported the SL effects of single-tone sequences to be right-hemisphere dependent (Roser et al., 2011; Danckert et al., 2012; Shaqiri and Anderson, 2013). The amplitude difference could not be retained after the statistical regularities of the two sequences were exchanged in the last third of each sequence, although the finding that the amplitude in the right hemisphere was lower than that in the left was retained. This may imply that learning effects cannot be retained when sequential regulations in the low and high voices are exchanged. A previous study has suggested that explicit knowledge cannot be transformed into implicit knowledge through practice (Krashen, 1982; Hulstijn, 2002). In contrast, other researchers have claimed that implicit and explicit knowledge can interact with each other (DeKeyser, 2003, 2007; De Jong, 2005; Ellis, 2005, 2009). The present study may imply that implicit and explicit learning can interact with each other, but only barely.

SL is reflected in the early component of P1 (Paraskevopoulos et al., 2012; Daikoku et al., 2016, 2017c) as well as in the

FIGURE 4 | Mean peak amplitudes (upper) and the latencies (lower) of P1m. Green bars: responses to chords consisting of two rare tones in attended and ignored sequences; pink bars: responses to chords consisting of a rare tone in attended sequence and a frequent tone in ignored sequence; blue bars: responses to chords consisting of a frequent tone in attended sequence and a rare tone in ignored sequence; black bars: responses to chords consisting of two frequent tones in attended and ignored sequences. Asterisks indicate significant differences ($p < 0.05$, Bonferroni-corrected).

late components such as N1, mismatch negativity (MMN), P2, and N400 (Abla et al., 2008; Furl et al., 2011; Daikoku et al., 2014, 2015; Koelsch et al., 2016). It is, however, considered that the SL effect relationship with P1 involves music expertise and specialized training experience (Boutros et al., 1995; Boutros and Belger, 1999; Kisley et al., 2004; Kizkin et al., 2006; Wang et al., 2009). According to a previous study (Adler et al., 1982), earlier auditory responses such as P1 were attributed to parallel thalamo-cortical connections and superior temporal gyrus. Thus, the findings of P1 in the present study can be interpreted as a prediction error suppressed by top-down predictions (Friston, 2005). Further studies are needed to reveal the role of P1 in SL.

Previous studies have suggested that the brain regions and activation patterns engaged during attentional and nonattentional learning might be partially distinct (Curran and Keele, 1993; Rauch et al., 1995; Reber and Squire, 1998; Jimenez and Castor, 1999; Poldrack et al., 2001; Aizenstein et al., 2004; Paradis, 2004; Destrebecqz et al., 2005; Daikoku and Yumoto, 2017). In our recent study, the SL of two simultaneous sequences was facilitated by paying attention to only one sequence and ignoring the other (Daikoku and Yumoto, 2017). This suggests that there is a partially distinct neural basis of attentional and nonattentional SL. In other words, biased attention might be an essential strategy in situations where the learner is exposed to multiple streams of information simultaneously. In this study, we exchanged the Markov model between attentional and nonattentional sequences in the last third of the sequences. We also revealed that the SL of two simultaneous auditory sequences might be right-hemisphere dependent. Learning effects cannot be retained when the tone sequence to which the subject is attending is changed during listening. These results suggest that attentional and nonattentional learning scarcely interact with each other and that there may be a specific cognitive system for nonattentional learning that is independent of attentional learning. As we could not demonstrate a neurological dichotomy between nonattentional and attentional SL due to the methodological limitations of the present study, further studies are needed to examine distinct or common neural mechanisms between attentional and nonattentional learning.

AUTHOR CONTRIBUTIONS

The experimental paradigms of the present study were considered by both of the authors. TD made the paradigms, recruited the participants, and collected the data; analyzed all of the data. MY proposed methodologies of MEG/behavioral data analyses and figures in the manuscript. Both of the authors discussed how the results could neurophysiologically and psychologically be interpreted. TD prepared the figures,

and wrote the main manuscript text. Then, both of the authors reviewed and revised the manuscript.

FUNDING

REFERENCES

Abla, D., Katahira, K., and Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *J. Cogn. Neurosci.* 20, 952–964. doi: 10.1162/jocn.2008.20058

Adler, L. E., Pachtman, E., Franks, R. D., Pecevich, M., Waldo, M. C., and Freedman, R. (1982). Neurophysiological evidence for a defect in neuronal mechanisms involved in sensory gating in schizophrenia. *Biol. Psychiatry* 17, 639–654.

Aizenstein, H. J., Stenger, V. A., Cochran, J., Clark, K., Johnson, M., Nebes, R. D., et al. (2004). Regional brain activation during concurrent implicit and explicit sequence learning. *Cereb. Cortex* 14, 199–208. doi: 10.1093/cercor/bhg119

Batterink, L. J., Reber, P. J., Neville, H. J., and Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *J. Mem. Lang.* 83, 62–78. doi: 10.1016/j.jml.2015.04.004

Boutros, N. N., and Belger, A. (1999). Midlatency evoked potentials attenuation and augmentation reflect different aspects of sensory gating. *Biol. Psychiatry* 45, 917–922. doi: 10.1016/s0006-3223(98)00253-4

Boutros, N. N., Torello, M. W., Barker, B. A., Tueting, P. A., Wu, S. C., and Nasrallah, H. A. (1995). The P50 evoked potential component and mismatch detection in normal volunteers: implications for the study of sensory gating. *Psychiatry Res.* 57, 83–88. doi: 10.1016/0165-1781(95)02637-c

Cleeremans, A., Destrebecqz, A., and Boyer, M. (1998). Implicit learning: news from the front. *Trends Cogn. Sci.* 2, 406–416. doi: 10.1016/s1364-6613(98)01232-7

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., et al. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48, 541–553. doi: 10.1016/j.neuroimage. 2009.06.069

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., and Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Res.* 1123, 168–178. doi: 10.1016/j. brainres.2006.09.046

Curran, T., and Keele, S. W. (1993). Attentional and non-attentional forms of sequence learning. *J. Exp. Psychol. Learn. Mem. Cogn.* 19, 189–202. doi: 10.1037/0278-7393.19.1.189

Daikoku, T. (2018a). Time-course variation of statistics embedded in music: corpus study on implicit learning and knowledge. *PLoS One* 13:e0196493. doi: 10.1371/journal.pone.0196493

Daikoku, T. (2018b). Neurophysiological markers of statistical learning in music and language: hierarchy, entropy, and uncertainty. *Brain Sci.* 8:114. doi: 10.3390/brainsci8060114

Daikoku, T. (2018c). Entropy, uncertainty, and the depth of implicit knowledge on musical creativity: Computational study of improvisation in melody and rhythm. *Front. Comput. Neurosci.* 12:97. doi: 10.3389/fncom.2018.00097

Daikoku, T. (2018d). Musical creativity and depth of implicit knowledge: Spectral and temporal individualities in improvisation. *Front. Comput. Neurosci.* 12:89. doi: 10.3389/fncom.2018.00089

Daikoku, T., Okano, T., and Yumoto, M. (2017a). "Relative difficulty of auditory statistical learning based on tone transition diversity modulates chunk length in the learning strategy," in *Proceedings of the Biomagnetic* (Sendai, Japan), 75.

Daikoku, T., Takahashi, Y., Futagami, H., Tarumoto, N., and Yasuda, H. (2017b). Physical fitness modulates incidental but not intentional statistical learning of simultaneous auditory sequences during concurrent physical exercise. *Neurol. Res.* 39, 107–116. doi: 10.1080/01616412.2016.1273571

Daikoku, T., Yatomi, Y., and Yumoto, M. (2017c). Statistical learning of an auditory sequence and reorganization of acquired knowledge: a time course of word segmentation and ordering. *Neuropsychologia* 95, 1–10. doi: 10.1016/j. neuropsychologia.2016.12.006

Daikoku, T., Takahashi, Y., Tarumoto, N., and Yasuda, H. (2018). Auditory statistical learning during concurrent physical exercise and the tolerance for pitch, tempo, and rhythm changes. *Motor Control* 22, 233–244. doi: 10.1123/mc.2017-0006

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

Daikoku, T., Yatomi, Y., and Yumoto, M. (2015). Statistical learning of music- and language-like sequences and tolerance for spectral shifts. *Neurobiol. Learn. Mem.* 118, 8–19. doi: 10.1016/j.nlm.2014.11.001

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

Daikoku, T., and Yumoto, M. (2017). Single, but not dual, attention facilitates statistical learning of two concurrent auditory sequences. *Sci. Rep.* 7:10108. doi: 10.1038/s41598-017-10476-x

Danckert, J., Stöttinger, E., Quehl, N., and Anderson, B. (2012). Right hemisphere brain damage impairs strategy updating. *Cereb. Cortex* 22, 2745–2760. doi: 10.1093/cercor/bhr351

De Diego Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., and Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One* 2:e1175. doi: 10.1371/journal.pone. 0001175

De Jong, N. (2005). *Learning Second Language Grammar by Listening (Unpublished Ph.D. thesis).* Netherlands: Graduate School of Linguistics.

DeKeyser, R. M. (2003). "Implicit and explicit learning," in *Handbook of Second Language Acquisition*, eds C. J. Doughty and M. H. Long (Oxford, MA: Blackwell), 313–348.

DeKeyser, R. M. (2007). *Practicing in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology.* Cambridge: Cambridge University Press.

Destrebecqz, A., Peigneux, P., Laureys, S., Degueldre, C., Del Fiore, G., Aerts, J., et al. (2005). The neural correlates of implicit and explicit sequence learning: interacting networks revealed by the process dissociation procedure. *Learn. Mem.* 12, 480–490. doi: 10.1101/lm.95605

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: a psychometric study. *Stud. Second Lang. Acquis.* 27, 141–172. doi: 10.1017/S0272263105050096

Ellis, R. (2009). "Implicit and explicit learning, knowledge and instruction," in *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*, eds R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip and H. Reinders (Bristol, England: Multilingual Matters), 3–25.

François, C., Chobert, J., Besson, M., and Schön, D. (2013). Music training for the development of speech segmentation. *Cereb. Cortex* 23, 2038–2043. doi: 10.1093/cercor/bhs180

François, C., Cunillera, T., Garcia, E., Laine, M., and Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia* 98, 56–67. doi: 10.1016/j.neuropsychologia.2016.10.006

François, C., and Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb. Cortex* 21, 2357–2365. doi: 10.1093/cercor/bhr022

François, C., Jaillet, F., Takerkart, S., and Schön, D. (2014). Faster sound streamsegmentation inmusicians than in nonmusicians. *PLoS One* 9:e101340. doi: 10.1371/journal.pone.0101340

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Furl, N., Kumar, S., Alter, K., Durrant, S., Shawe-Taylor, J., and Griffiths, T. D. (2011). Neural prediction of higher-order auditory sequence statistics. *Neuroimage* 54, 2267–2277. doi: 10.1016/j.neuroimage.2010.10.038

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Lang. Res.* 18, 193–223. doi: 10.1191/0267658302sr207oa

Jimenez, L., and Castor, M. (1999). Which attention is needed for implicit sequence learning? *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 236–259. doi: 10.1037/0278-7393.25.1.236

Kisley, M. A., Noecker, T. L., and Guinther, P. M. (2004). Comparison of sensory gating to mismatch negativity and self-reported perceptual phenomena in healthy adults. *Psychophysiology* 41, 604–612. doi: 10.1111/j.1469-8986.2004. 00191.x

Kizkin, S., Karlidag, R., Ozcan, C., and Ozisik, H. I. (2006). Reduced P50 auditory sensory gating response in professional musicians. *Brain Cogn.* 61, 249–254. doi: 10.1016/j.bandc.2006.01.006

Koelsch, S., Busch, T., Jentschke, S., and Rohrmeier, M. (2016). Under the hood of statistical learning: a statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Sci. Rep.* 6:19741. doi: 10.1038/srep19741

Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition.* Oxford: Pergamon.

Markov, A. A. (1971). "Extension of the limit theorems of probability theory to a sum of variables connected in a chain," in reprinted in Appendix B of: R. Howard. D. *Dynamic Probabilistic Systems, volume 1 Markov Chains*, ed. R. Howard (New york, NY: John Wiley and Sons), 552–577.

Moldwin, T., Schwartz, O., and Sussman, E. S. (2017). Statistical learning of melodic patterns influences the brain's response to wrong notes. *J. Cogn. Neurosci.* 29, 2114–2122. doi: 10.1162/jocn_a_01181

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

Paradis, M. (2004). *A Neurolinguistic Theory of Bilingualism.* Amsterdam: John Benjamins.

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., and Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: an MEG study. *Neuropsychologia* 50, 341–349. doi: 10.1016/j.neuropsychologia.2011.12.007

Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends Cogn. Sci.* 10, 233–238. doi: 10.1016/j. tics.2006.03.006

Poldrack, R. A., Clark, J., PareÂ-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., et al. (2001). Interactivememory systems in the human brain. *Nature* 414, 546–550. doi: 10.1038/35107080

Rauch, S. L., Savage, C. R., Brown, H. D., Curran, T., Alpert, N. M., Kendrick, A., et al. (1995). A PET investigation of implicit and explicit sequence learning. *Hum. Brain Mapp.* 3, 271–286. doi: 10.1002/hbm. 460030403

Reber, A. S. (1989). Implicit learning and tacit knowledge. *J. Exp. Psychol. Gen.* 118, 219–235. doi: 10.1037/0096-3445.118.3.219

Roser, M. E., Fiser, J., Aslin, R. N., and Gazzaniga, M. S. (2011). Right hemisphere dominance in visual statistical learning. *J. Cogn. Neurosci.* 23, 1088–1099. doi: 10.1162/jocn.2010.21508

Reber, P. J., and Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *J. Cogn. Neurosci.* 10, 248–263. doi: 10.1162/089892998562681

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928. doi: 10.1126/science.274.5294.1926

Sanders, L. D., Newport, E. L., and Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nat. Neurosci.* 5, 700–703. doi: 10.1038/nn873

Schön, D., and François, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Front. Psychol.* 2:167. doi: 10.3389/fpsyg.2011. 00167

Shaqiri, A., and Anderson, B. (2013). Priming and statistical learning in right brain damaged patients. *Neuropsychologia* 51, 2526–2533. doi: 10.1016/j. neuropsychologia.2013.09.024

Skoe, E., Krizman, J., Spitzer, E., and Kraus, N. (2015). Prior experience biases subcortical sensitivity to sound patterns. *J. Cogn. Neurosci.* 27, 124–140. doi: 10.1162/jocn_a_00691

Wang, W., Staffaroni, L., Reid, E., Steinschneider, M., and Sussman, E. (2009). Effects of musical training on sound pattern processing in highschool students. *Int. J. Pediatr. Otorhinolaryngol.* 73, 751–755. doi: 10.1016/j.ijporl.2009. 02.003

Yumoto, M., and Daikoku, T. (2016). "IV Auditory system. 5 basic function," in *Clinical Applications of Magnetoencephalography*, eds S. Tobimatsu and R. Kakigi (Berlin/Heidelberg, Germany: Springer), 97–112.

Yumoto, M., and Daikoku, T. (2018). Neurophysiological study of statistical learning of auditory stimulation sequences. *J. Cogn. Neurosci.* 20, 38–43.