



ELSEVIER

Contents lists available at ScienceDirect

Neuropsychologia

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

Implicit and explicit statistical learning of tone sequences across spectral shifts



Tatsuya Daikoku, Yutaka Yatomi, Masato Yumoto*

Department of Clinical Laboratory, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan

ARTICLE INFO

Article history:

Received 13 March 2014

Received in revised form

21 August 2014

Accepted 22 August 2014

Available online 2 September 2014

Keywords:

Implicit and explicit learning

Statistical learning

Markov process

Relative pitch

Magnetoencephalography

ABSTRACT

We investigated how the statistical learning of auditory sequences is reflected in neuromagnetic responses in implicit and explicit learning conditions. Complex tones with fundamental frequencies (F0s) in a five-tone equal temperament were generated by a formant synthesizer. The tones were subsequently ordered with the constraint that the probability of the forthcoming tone was statistically defined (80% for one tone; 5% for the other four) by the latest two successive tones (second-order Markov chains). The tone sequence consisted of 500 tones and 250 successive tones with a relative shift of F0s based on the same Markov transitional matrix. In explicit and implicit learning conditions, neuromagnetic responses to the tone sequence were recorded from fourteen right-handed participants. The temporal profiles of the N1m responses to the tones with higher and lower transitional probabilities were compared. In the explicit learning condition, the N1m responses to tones with higher transitional probability were significantly decreased compared with responses to tones with lower transitional probability in the latter half of the 500-tone sequence. Furthermore, this difference was retained even after the F0s were relatively shifted. In the implicit learning condition, N1m responses to tones with higher transitional probability were significantly decreased only for the 250 tones following the relative shift of F0s. The delayed detection of learning effects across the sound-spectral shift in the implicit condition may imply that learning may progress earlier in explicit learning conditions than in implicit learning conditions. The finding that the learning effects were retained across spectral shifts regardless of the learning modality indicates that relative pitch processing may be an essential ability for humans.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

In recent decades, the relationship between the learner's awareness of learning and the effect of learning has been heavily researched in cognitive science. According to previous studies, the learning system is divided into two major mechanisms: implicit learning and explicit learning. Implicit learning is generally defined as learning without the learner's being aware of having learned anything. In contrast, explicit learning is defined as learning with the learner's being aware of having learned (Ellis, 2005, 2009; Reber, 1989).

1.1. Implicit and explicit knowledge

Implicit knowledge was first reported by researchers who had been studying implicit consciousness (Schacter, 1987). Implicit knowledge is considered a type of knowledge that learners cannot verbalize, despite evidence of having learned as detected from the

learner's behavioral responses (Ellis, 2005, 2009; Schacter, 1987). In contrast, explicit knowledge is knowledge that learners can verbalize and consciously remember (Hulstijn, 2002, 2005; Jacoby, 1983). Furthermore, earlier studies have suggested that explicit knowledge is declarative, while implicit knowledge is procedural (Clark and Squire, 1998; De Jong, 2005; Ellis, 2005, 2009; Paradis, 2004; Ullman, 2001). According to neurophysiological studies of patients with cognitive impairment, the implicit knowledge system may be impaired in cases of aphasia and Parkinson's disease, whereas the explicit knowledge system may be damaged in cases of amnesia, Alzheimer's disease and Williams' syndrome (Clark & Squire, 1998; Paradis, 2004; Squire, Knowlton, & Musen 1993; Ullman, 2001), thus suggesting that different regions of the brain may retain different types of knowledge. Accordingly, these findings support the dichotomy between implicit and explicit knowledge.

1.2. Implicit and explicit learning

With respect to learning, it is difficult to distinguish between implicit and explicit processes (DeKeyser, 2003; De Jong, 2005; Ellis, 2005, 2009). Schmidt argued that the "noticing", which

* Corresponding author. Tel.: +81 3 3815 5411; fax: +81 3 5689 0495.

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

typically involves some degree of awareness, is essential for both explicit and implicit learning. Based on this argument, there is no such thing as unconscious implicit learning (Schmidt, 1994, 2001). Schmidt also suggested that implicit learning might be a type of learning that occurs without any metacognition or higher-order awareness (i.e., awareness of learning). Furthermore, many studies have indicated methodological limitations with respect to assessing the degree of the learner's awareness of learning (DeKeyser, 2003; De Jong, 2005; Ellis, 2005, 2009). However, there are also studies that explained the differences between explicit and implicit learning processes. First, these two types of learning processes differ with respect to the neural regions that are activated during learning (Destrebecqz et al., 2005; Paradis, 2004). Second, explicit learning tends to favor short-term knowledge (i.e., working memory), whereas implicit learning tends to favor long-term knowledge (Ellis, 2009). Furthermore, differences in the effects of the two learning processes have also been reported in many earlier studies. While most studies suggested that explicit learning is more effective than implicit learning, these studies measured the learning effects using behavioral tests that may favor explicit learning such as short-term learning. Other studies suggested that the learning effects depend on the learner's prior knowledge (Francois & Schön, 2011; Gass, Svetics, & Lemelin, 2003). Thus, we cannot precisely determine which learning process is more effective (DeKeyser, 2003; Ellis, 2009). Interestingly, some earlier studies claimed that knowledge of music and language is acquired largely by implicit learning (Ettlinger, Margulis, & Wong, 2011; Francois & Schön, 2011; Krashen, 1982). Reber (1989) argued that implicit learning is the fundamental process that lies at the heart of the adaptive behavioral repertoire of every complex organism, while DeKeyser claimed that implicit learning is essential for adaptation to the environment (Reber, 1993). He said that children acquire language better than adults in terms of "ultimate learning achievement", although adults acquire a language faster than children do. This notion may be supported by results of a study that indicate explicit knowledge is more likely to lead to errors in judgment than is implicit knowledge (Dienes & Scott, 2005) as children tend to learn a language implicitly while adults tend to learn a language explicitly. Furthermore, it is argued that children might be the best learners in all generations if they are given ample time to learn (DeKeyser, 2003).

1.3. Relationship between learning and knowledge

Many early studies have investigated the relationship between learning and knowledge. For example, Hulstijn and Krashen suggested that because implicit and explicit knowledge are acquired through different mechanisms, explicit knowledge cannot be transformed into implicit knowledge through practice. This premise is called the "noninterface" position (Hulstijn, 2002; Krashen, 1982). Other researchers, however, have claimed that implicit and explicit knowledge can interact with one another, a concept referred to as the "interface" position (DeKeyser, 2003, 2007; De Jong, 2005; Ellis, 2005, 2009). Researchers who supported either the strong or weak interface position claimed that not only can explicit knowledge be derived from implicit knowledge but also that explicit knowledge can be converted into implicit knowledge through practice (DeKeyser, 2003, 2007; Ellis, 2005, 2009). Thus, learning conditions (explicit/implicit learning) do not restrict the type of knowledge that learners acquire.

1.4. Relative pitch processing

It has recently been clarified that a learner's brain activity may depend on the learner's prior knowledge (Furl et al., 2011; Pantev et al., 1998; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev,

2012; Yumoto et al., 2005). This suggests that if the fundamentals that learners have already acquired are used in an experiment, the level of learning achievement cannot be balanced across all participants. Therefore, in this study, we devised original rules that the participants were unlikely to have previously encountered. This allowed us to minimize possible biases induced by differences in the prior knowledge of individual participants. However, the learning regularities that people have never experienced and therefore cannot predict are just the first step in learning. In most learning activities of healthy humans, newly encountered regularities are considered in relation to regularities already acquired. An analogous notion is found in the information theory (Olshausen & Field, 2004). If we separately recognize all information we receive as entirely different information, we must cope with a vast amount of information, a process that is systematically redundant. On the other hand, if we recognize many types of information by correlating information that we have already learned, we do not have to integrate all received information, thereby sparing memory capacity in our brains. This process is systematically efficient. One example of this efficient processing is relative pitch. Listeners can easily recognize transposed melodies if they have previously listened to the original melodies. Even infants can possess relative pitch (Plantinga & Trainor, 2005; Trehub, 2001; Trehub & Hannon, 2006) as statistical learning may implicitly promote the acquisition of relative pitch in infants (Saffran, 2003). In this study, we investigated whether adult participants could statistically learn higher-order structures with tone sequences in both implicit and explicit conditions. To clarify further the detailed mechanisms of implicit and explicit learning, it is important to reveal how relative pitch processing is implemented in both implicit and explicit learning in humans.

1.5. Event-related responses as an index of statistical learning

The mismatch negativity (MMN) is a component of the event-related responses to deviants in sound sequences, peaking at approximately 100–250 ms after the onset of the deviant stimulus (Haenschel, Vernon, Dwivedi, Gruzelić, & Baldeweg, 2005; Näätänen, Paavilainen, Rinne, & Alho, 2007; Shestakova et al., 2002). In previous studies on statistical learning using word segmentation paradigms, the amplitudes of the N1 responses to tones with lower transitional probability were increased compared with responses to tones with higher transitional probability (Abla, Katahira, & Okanoya, 2008; Paraskevopoulos et al., 2012). Thus, the N1 responses have been used as an index of statistical learning of auditory sequences. Furthermore, Paraskevopoulos et al. suggested that modulation of the neural responses in the latency range of N1m during statistical learning could be interpreted as MMN.

In several fields of study, such as natural language processing (Poon & Domingos, 2007, 2008; Singla & Domingos, 2006), music perception and statistical learning (Richardson & Domingos, 2006), the Markov chain has often been used as a model of the artificial grammar of language and music. The Markov chain, which was first reported by Andrey Markov (1971), is a mathematical system in which the probability of the forthcoming state is statistically defined only by the latest state. The word segmentation paradigm is also considered a specific form of the Markov chain. The use of the Markov chains embedded in tone sequences allows us to verify the mechanism of statistical learning in the acquisition of language and music. In the present study, using tone sequences based on second-order Markov chains that have a more general structure than found in the word segmentation task, we investigated the hypotheses that not only statistical learning but also the relative pitch processing of auditory sequences based on

statistical regularities are reflected in the N1m (the magnetic counterpart of the N1 potential) responses in implicit and explicit conditions.

2. Methods

2.1. Participants

Fourteen right-handed (Edinburgh handedness questionnaires; laterality quotient ranged from 57.9 to 100) (Oldfield, 1971) healthy Japanese participants with no history of neurological or audiological disorders were included (8 males, 6 females; age range: 24–36 years). None of the participants possessed absolute pitch. This study was approved by the Ethics Committee of The University of Tokyo. All participants were informed of the purpose of the study and ensured that their personal data would be protected. All participants gave written informed consent.

2.2. Stimuli

2.2.1. Tones

Using a Klatt-type cascade synthesizer (Klatt, 1980) HLSyn (Sensimetrics Corporation, Malden, MA, USA), we generated complex tones with fundamental frequencies (F_0) in a five-tone equal temperament ($F_0 = 100 \times 2^{(n-1)/5}$ Hz, $n = 1-7$; 100, 115, 132, 152, 174, 200, and 230 Hz) and the first and second formant frequencies of 430 and 1000 Hz, respectively. Only the F_0 s were variable. All other parameters were constant: duration 400 ms including 10 and 200 ms rise and decay time, respectively, and binaural presentation with an intensity of 80 dB SPL.

2.2.2. Sequences

The order of tones was defined by a second-order Markov process with the constraint that the probability of a forthcoming tone was statistically defined (80% for one tone; 5% for the other four tones) by the latest two successive tones. Fig. 1 shows the two transitional matrices of the Markov chains used in the present study. Each pair of two tones (row) could be followed by one of the five subsequent tones (columns) with a probability of either 80% or 5%. As the two Markov chains shown in Fig. 1 were adopted in the two experiment conditions, implicit and explicit learning conditions, the particular transitional patterns did not interfere with learning in the adjacent experiment. Furthermore, the order of the adoption was counterbalanced across the participants.

The sequence consisted of 750 successive tones and was presented with a stimulus onset asynchrony (SOA) of 600 ms. The sequence comprised three portions (first, middle, and last) with 250 tones each (Fig. 2). The first and middle portions consisted of the tones with the five F_0 s ($F_0 = 100 \times 2^{(n-1)/5}$ Hz, $n = 1-5$). The last portion consisted of tones with F_0 s ($F_0 = 100 \times 2^{(n-1)/5}$ Hz) shifted from $n = 1-5$ to $n = 3-7$, but the transitional patterns were not changed (Fig. 1b). We recorded magnetoencephalographic (MEG) signals from the participants as they listened to the tone sequences.

2.3. Experimental protocol

First, the experiment in the implicit learning condition was conducted. After approximately two minutes that included behavioral tests, the experiment in the explicit learning condition was conducted. The two experiments had to be performed in this order because the regularity of tone sequences could not be disclosed to the participants until the end of the experiment in the implicit condition.

2.3.1. Experiment 1 (implicit learning condition)

The participants were shown a silent movie during the auditory stimulus presentation and asked a set of questions about the detailed contents of the movie (behavioral test) after the MEG measurement. The silent movie was one that none of the participants had previously seen, and the questions were ones that could not be answered unless the participants carefully attended to the entire movie (e.g., the number of characters). Before the experiments, the participants were provided the questions as deductive instruction, and they were told that they would be tested on the movie after the MEG measurement as intentional instruction. Deductive and intentional instructions have been pre-eminent in differentiating between explicit and implicit learning conditions in recent studies on learning (Alanen, 1995; DeKeyser, 2003; Ellis, 1993, 2009; Leow, 1998; Michas & Berry, 1994; Robinson, 1996; Rosa & O'Neill, 1999). Under these conditions, we can assume that they paid more attention to the silent movie than to the tone sequences. This behavioral test was completed within two minutes for each participant.

2.3.2. Experiment 2 (explicit learning condition)

In all the tone sequences, a one-second silent period was pseudo-randomly inserted (i.e., SOA 1.6 s) within every set of thirty successive tones. During the

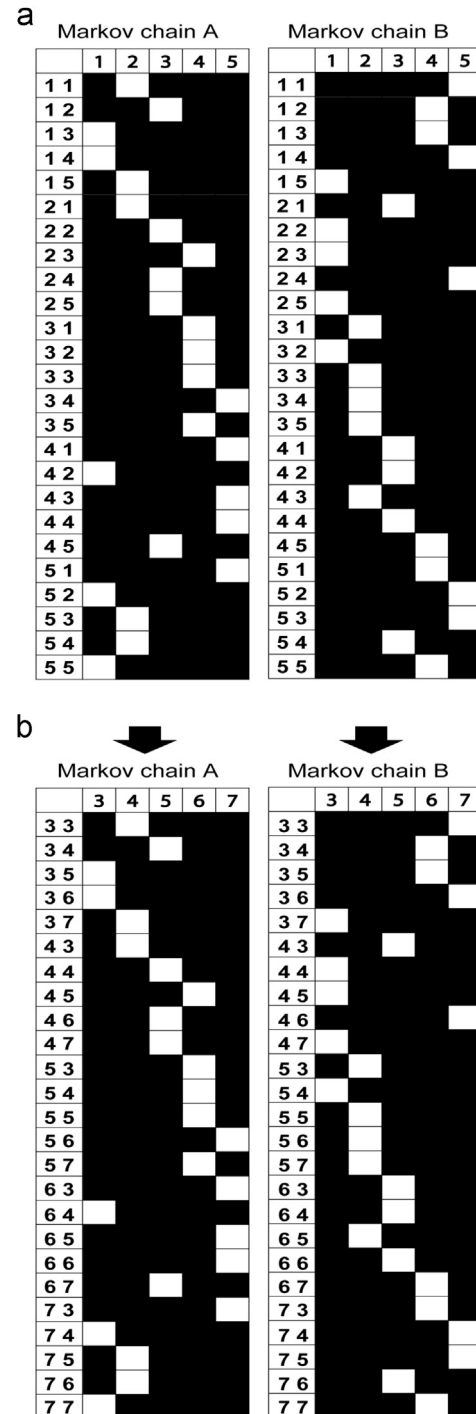


Fig. 1. The transitional matrices of the Markov chains used in the present study. *Rows*: the two tones most recently presented. *Columns*: the next tone that may appear. Each pair of tones (rows) could precede one of the five subsequent tones (columns) with an 80% (white cells) or 5% (black cells) probability. The transitional matrices used in the first and middle portions (a) and those used in the last portions (b).

experiment, the participants were instructed to raise their right hand at every silent period to confirm that they continued to pay attention to the tone sequences. Before the experiment, participants were told that the tones were sequenced on the basis of transitional probabilities as a deductive instruction and that they would be tested on the tone sequences after the experiments as intentional instruction. With these conditions, we can assume that they continued to pay attention to the tone sequences. After the MEG measurement, the participants were presented 10 tone series with 8 tones, half of which were sequenced by the same Markov

process. The participants were asked whether each tone series sounded familiar to them. If a participant had sufficiently learned the transitional probabilities of the sequences, he or she should have been able to correctly answer this question. This behavioral test was completed within two minutes for each participant.

2.4. Measurement

The auditory stimuli were sequenced using the STIM2 system (Compumedics Neuroscan, El Paso, TX, USA) and were delivered binaurally to each participant's ears at 80 dB SPL through ER-3A earphones (Etymotic Research, Elk Grove Village, IL, USA). We recorded the MEG signals from the participants while listening to the tone sequences in a magnetically shielded room using a 306-channel neuromagnetometer system (Elekta Neuromag Oy, Helsinki, Finland) with 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 using a 1–200 Hz band-pass filter and then recorded at a sampling rate of 600 Hz. Epochs with artifacts of more than 3 pT/cm in any MEG channel were excluded from the analyses. Using the time-domain extension of signal space separation with a buffer length of 10 seconds and a correlation limit of .980, contamination from environmental magnetic noise was reduced (Taulu & Hari, 2009). The waveforms were filtered offline using a 2 to 40 Hz band-pass. The baseline for the waveforms in each MEG channel was defined by the mean amplitude between –100 and 0 ms.

2.5. Data analysis

The target of the analysis was the responses to the three tones that appeared in every portion of the sequences (i.e., $f_0 = 100 \times 2^{(n-1)/5}$ Hz, $n=3-5$) to exclude the

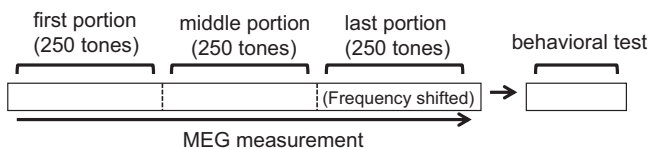


Fig. 2. Experimental design.

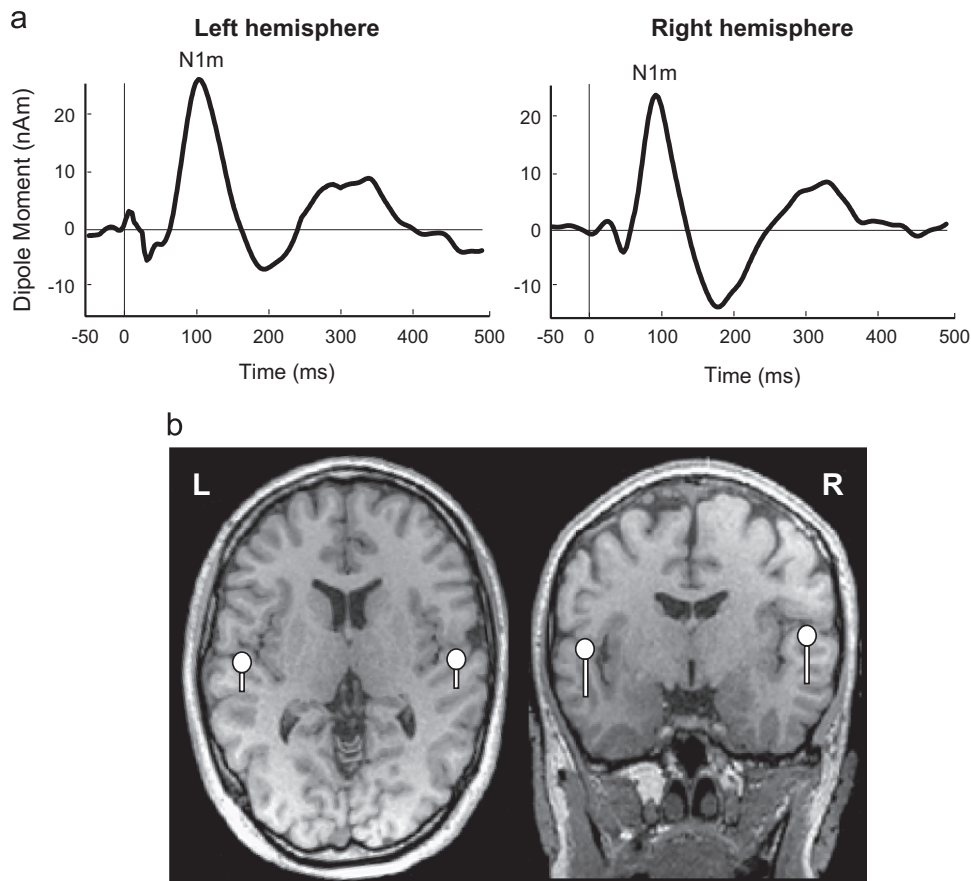


Fig. 3. (a) Source-strength waveforms for the N1m responses in a representative subject. (b) The ECD locations (dots) and orientations (bars) for the N1m responses superimposed on the MR images (1.5 T, TR/TE=2030/4.38 ms).

possibility that a difference detected in recorded responses could be attributed to the effect of physical differences in the presented tone ensembles. The noiseless MEG responses to the three tones were separately averaged according to the transitional probabilities from the beginning of each portion until the number of averages reached 30 (i.e., 10 times for each tone) for each portion. The sources of the N1m and P2m responses were modeled separately as a single equivalent current dipole (ECD) in each hemisphere. The ECDs for the N1m and P2m were separately calculated at the peak latencies with a goodness of fit above 80% using the same 66 temporal channels (44 gradiometers and 22 magnetometers) for each participant. The selected channel areas have been presented in a prior study (Yumoto et al., 2005). The source-strength waveforms for the N1m and P2m in each hemisphere were calculated using the ECDs (Fig. 3). Then, in each learning condition (implicit and explicit), we performed a 2 (hemisphere: right and left) \times 3 (portion: first, middle and last) \times 2 (probability: high and low) repeated-measures analysis of variance with the peak amplitude and the peak latency of the source-strength waveforms of the N1m and P2m. For further statistical analysis, we conducted post-hoc tests using the Bonferroni correction of the significance probability.

3. Results

3.1. Behavioral data

The average number of correct answers was far above the chance level in both the implicit learning condition (questions about the detailed contents of the movie) and the explicit learning condition (questions of whether each tone series sounded familiar) (percentages of correct answers, average \pm S.D.: implicit learning: $89.3 \pm 20.5\%$, explicit learning: $87.1 \pm 9.6\%$).

3.2. Magnetoencephalographic data

The averaged amplitudes and latencies of N1m and P2m responses are shown in Fig. 4.

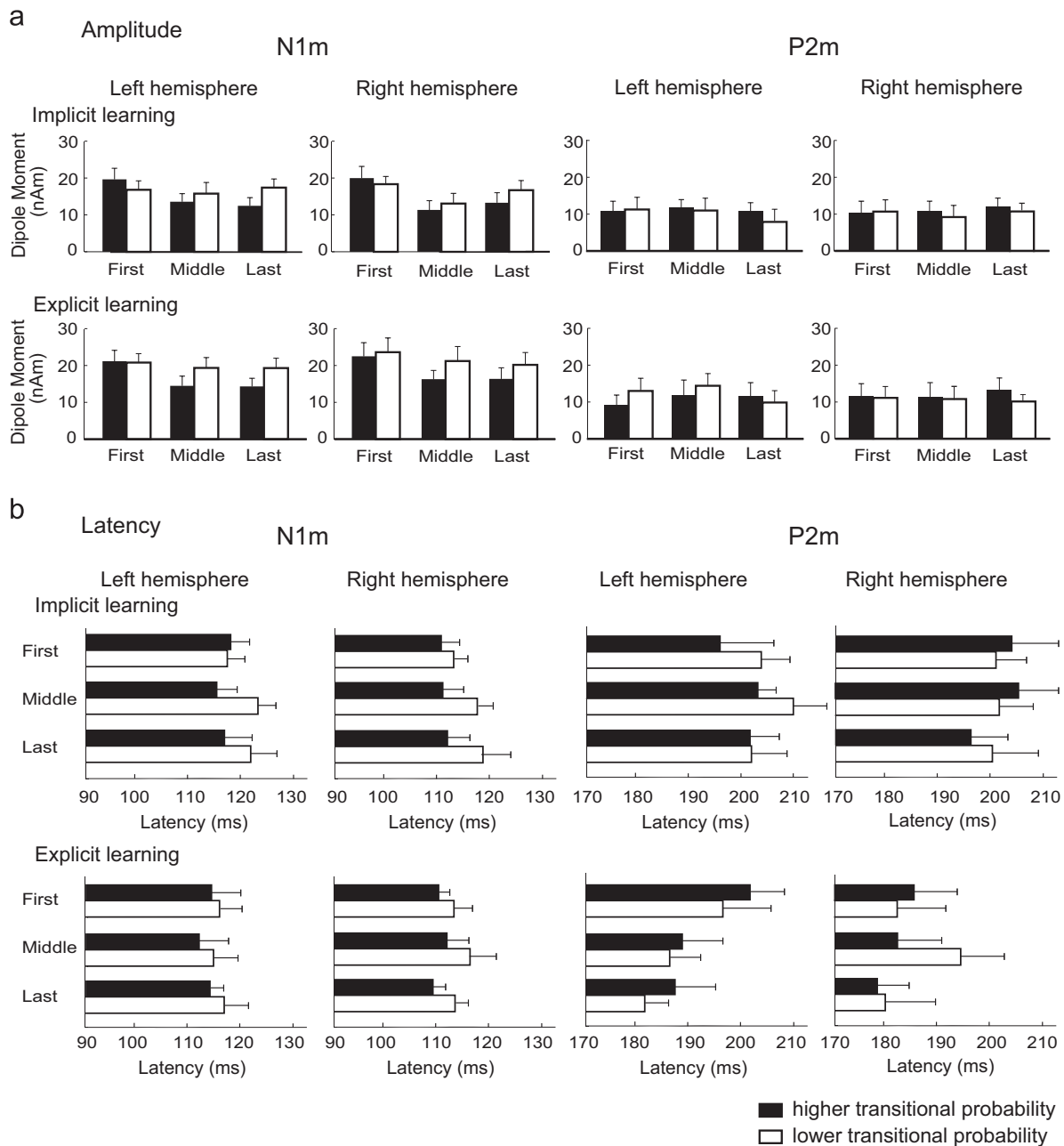


Fig. 4. The averaged peak amplitudes and latencies of the N1m and P2m responses in the implicit and explicit learning conditions ($N=14$). The responses in the left and right hemispheres are located on the left and right sides, respectively. The closed bars represent the responses to tones with higher transitional probability, and the open bars represent the responses to tones with lower transitional probability. The error bars indicate the standard error of the mean.

3.2.1. Experiment 1 (implicit learning condition)

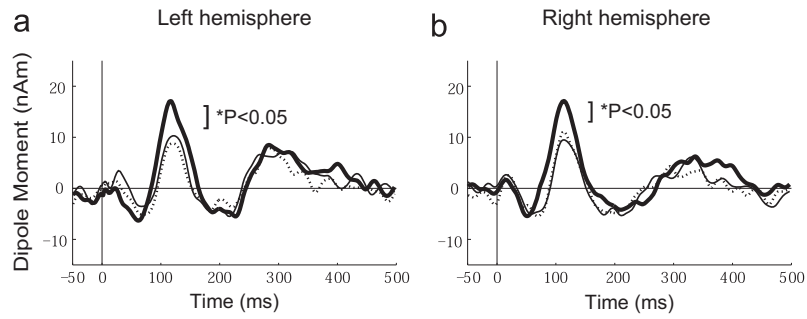
3.2.1.1. Amplitude. The main portion effect on the N1m peak amplitudes was significant ($F[2, 26]=9.90, p=.001$). The post-hoc test of the main portion effect revealed that the N1m peak amplitudes in the middle and last portions were significantly decreased compared with those in the first portion (middle portion: $p=.001$, last portion: $p=.046$). The portion–probability interaction of the N1m peak amplitudes was significant ($F[2, 26]=5.25, p=.012$). The post-hoc test of the portion–probability interaction revealed two significant results. First, the N1m peak amplitudes for tones with higher transitional probability in both the middle and last portions were significantly decreased compared with those in the first portion (middle portion: $p=.0001$, last portion: $p=.014$). However, no significant result was obtained for the N1m peak amplitudes for tones with lower

transitional probability (middle portion: $p=.13$, last portion: $p=1.0$) (Fig. 5). Second, with respect to the last portions, the N1m peak amplitudes for tones with higher transitional probability were significantly lower than for tones with lower transitional probability ($p=.005$) (Figs. 4 and 6).

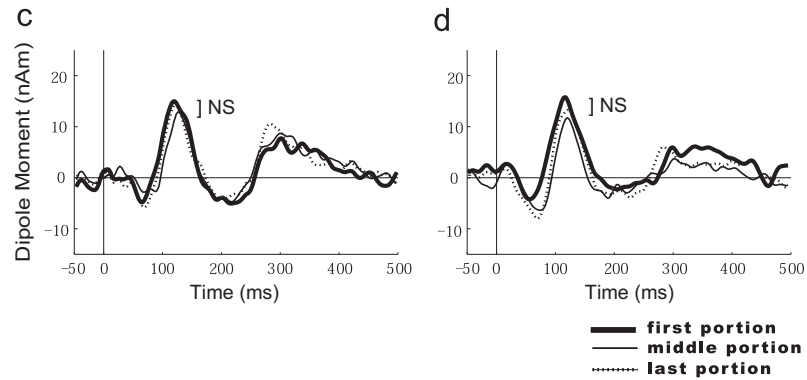
3.2.1.2. Latency. The main hemisphere effect and the main probability effect on the N1m peak latencies were significant (hemisphere: $F[1, 13]=4.93, p=.045$, probability: $F[1, 13]=5.86, p=.031$). The N1m peak latencies in the left hemisphere were significantly longer than those in the right hemisphere, while the N1m peak latencies for the tones with higher transitional probability were shorter than those for the tones with lower transitional probability. No other significant differences were detected in Experiment 1.

Implicit learning

Higher transitional probability

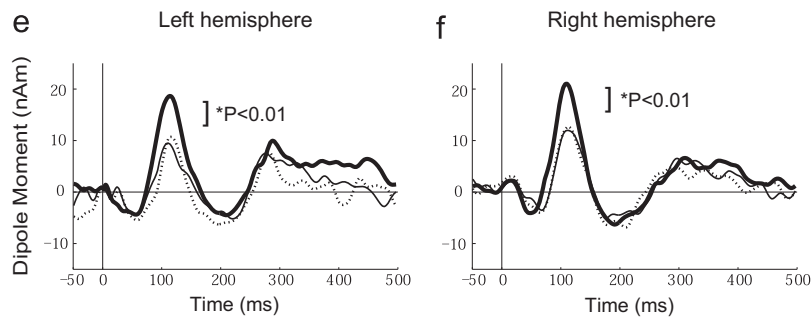


Lower transitional probability



Explicit learning

Higher transitional probability



Lower transitional probability

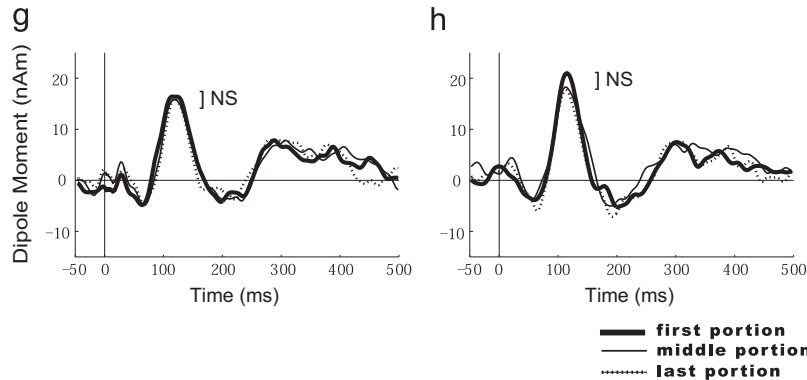


Fig. 5. Grand-averaged source-strength waveforms for the N1m responses ($N=14$). The responses to tones with higher transitional probability in the left (a) and right (b) hemispheres and to those with lower transitional probability in the left (c) and right (d) hemispheres in the implicit learning condition. The responses to tones with higher transitional probability in the left (e) and right (f) hemispheres and to those with lower transitional probability in the left (g) and right (h) hemispheres in the explicit learning condition. The thick lines represent the responses in the first portion, thin lines the responses in the middle portion, and dotted lines the responses in the last portion.

3.2.2. Experiment 2 (explicit learning condition)

3.2.2.1. Amplitude. The main portion effect and the main probability effect on the N1m peak amplitudes were significant (portion: $F[2, 26]=10.06, p=.001$, probability: $F[1, 13]=16.34, p=.001$). The post-hoc test of the main portion effect revealed that the N1m peak amplitudes in the middle and last portions were significantly decreased compared with those in the first portion (middle portion: $p=.001$, last portion: $p=.022$). Furthermore, the portion–probability interaction of the N1m peak amplitudes was significant ($F[2, 26]=4.80, p=.017$). The post-hoc test of the portion–probability interaction revealed two significant results. First, the N1m peak amplitudes for tones with higher transitional probability in both the middle and last portions were significantly decreased compared with those in the first portion (middle portion: $p=.004$, last portion: $p=.005$), but no significant result was obtained for the N1m peak amplitudes for tones with lower transitional probability (middle portion: $p=.15$, last portion: $p=.31$) (Fig. 5). Second, in both the middle and last portions, the N1m peak amplitudes for tones with higher transitional probability were significantly lower than those for tones with lower transitional probability (middle portion: $p=.003$, last portion: $p=.002$) (Figs. 4 and 6).

3.2.2.2. Latency. The main probability effect on the N1m peak latencies was significant ($F[1, 13]=9.05, p=.010$). The N1m peak latencies for the tones with higher transitional probability were shorter than those for the tones with lower transitional probability. No other significant differences were detected in Experiment 2.

4. Discussion

When learners have acquired statistical knowledge based on transitional probabilities of tone sequences, they can expect a tone that will follow certain preceding tones in the sequence.

With this expectation for upcoming tones, tones with higher transitional probability (i.e., more predictable tones) lead to decreases in neural responses compared with tones with lower transitional probability (i.e., less predictable tones). In contrast, the violation of this expectation based on the learned transitional probabilities leads to increased neural responses when compared with more predictable tones. Thus, differences in the amplitude of the learner's neural responses between tones with higher and lower

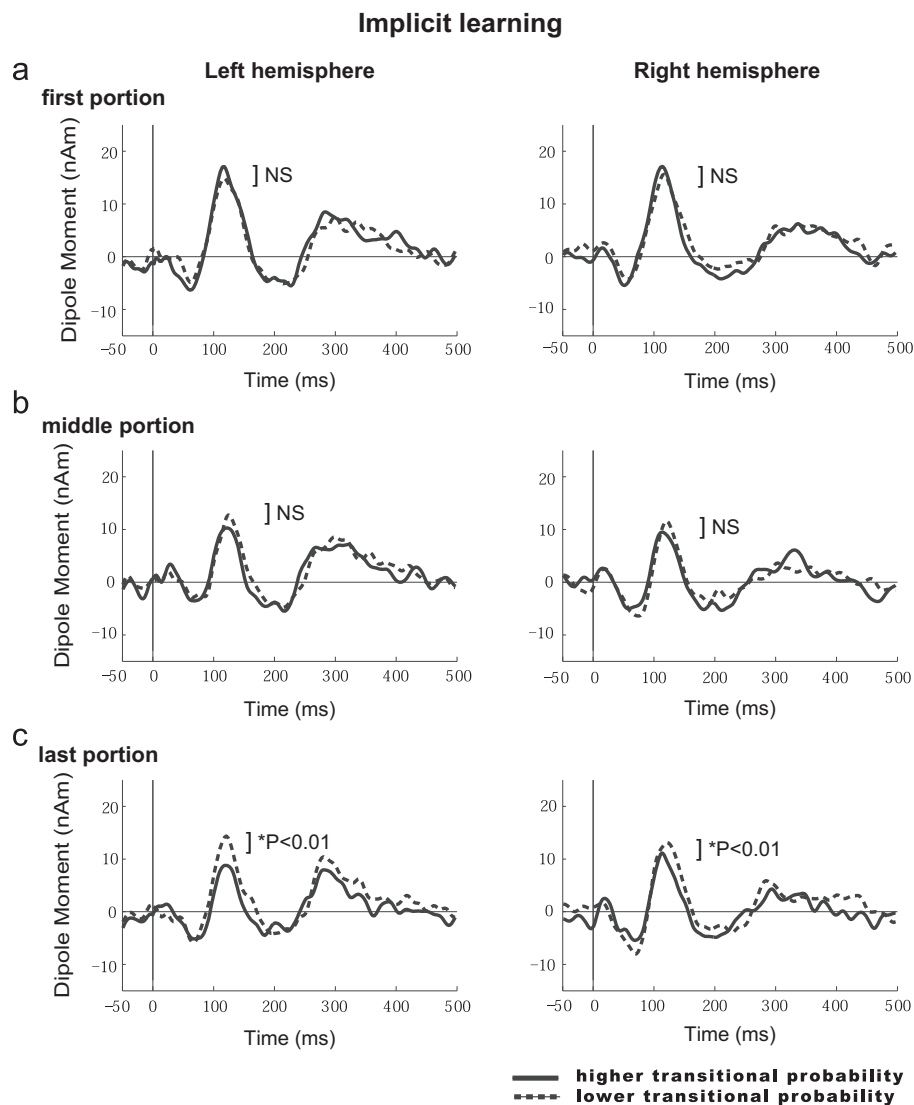


Fig. 6. Grand-averaged source-strength waveforms for the N1m responses ($N=14$). The first (a), middle (b), and last (c) portions of the implicit learning condition. The first (d), middle (e), and last (f) portions of the explicit learning condition. The responses in the left and right hemispheres are located on the left and right sides, respectively. The solid lines represent the responses to tones with higher transitional probability, the dashed lines represent the responses to tones with lower transitional probability.

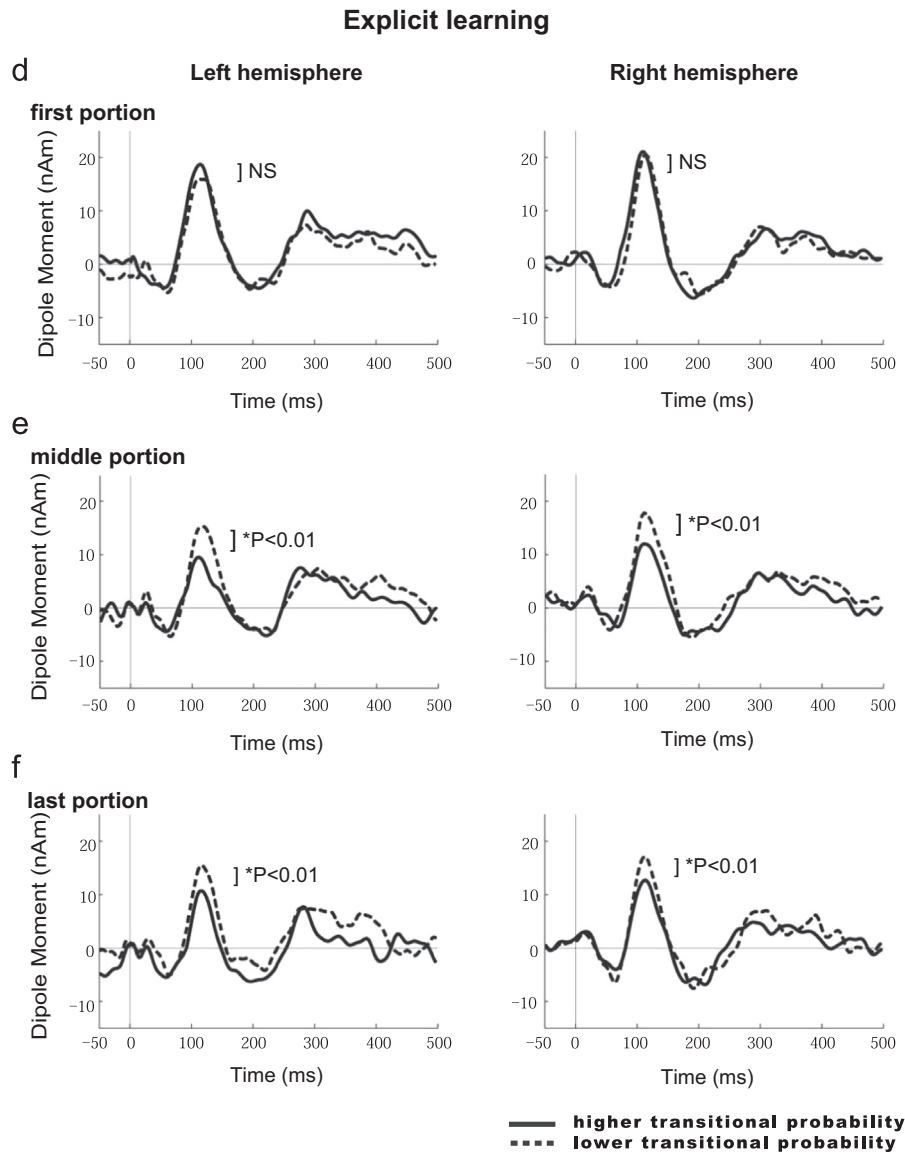


Fig. 6. (continued)

transitional probabilities occur after they have learned the transitional probabilities of tone sequences (Abla et al., 2008; Furl et al., 2011; Paraskevopoulos et al., 2012). However, neural responses to tones are gradually attenuated along with repetition due to the adaptation of auditory cortical neurons. The previous studies have reported that as repeated auditory stimulation leads to habituation of the activation of the auditory cortex, the N1m and P2m responses can be decreased (Kuriki, Ohta, & Koyama, 2007; Teismann et al., 2004). Therefore, learning and adaptation effects on amplitudes and latencies may be reflected in the neuromagnetic responses recorded in the present study. Furthermore, in the present study, the spectral frequencies of F0s were relatively shifted in the last portion of each sequence. If participants possess the ability to process relative pitch, the statistical learning effects should be retained even after the spectral frequencies are shifted.

In the explicit learning condition, amplitudes of responses to tones with lower transitional probabilities were significantly higher than those to tones with higher transitional probabilities, and responses to tones with higher transitional probability were significantly attenuated in the middle and last portions. These effects of statistical learning were retained even after the fundamental frequencies were shifted. Furthermore, a shortened latency

in the responses to tones with higher transitional probability was detected. In an earlier study, a similar tendency of latency shifts was observed, though it did not reach significance (Abla et al., 2008). Statistical learning may be reflected not only in the amplitude but also in the latency. With the expectation of upcoming tones, tones with higher transitional probability may lead to a shortening of the neural response latencies compared with tones with lower transitional probability.

In the implicit learning condition, responses to tones with higher transitional probability were significantly attenuated in the middle and last portions. In addition, these effects of statistical learning were retained even after fundamental frequencies were shifted. A shortening of neural response latencies was also detected. However, amplitudes of responses to tones with lower transitional probabilities were significantly higher than those to tones with higher transitional probabilities only in the last portion, although this tendency was also observed in the middle portion. In contrast, these significant effects were detected in both the middle and last portions in the explicit learning condition. The finding that learning effects were detected in the explicit learning condition earlier than in the implicit learning condition may imply that explicit learning is more effective than implicit learning, at least in

the early stage of knowledge acquisition. A number of earlier studies claimed that statistical learning and implicit learning are likely to favor long-term learning over short-term learning. Furthermore, most studies demonstrated that explicit learning could be more effective than implicit learning (Ellis, 2009). Therefore, it has been suggested that longer periods are necessary for statistical learning, particularly under implicit conditions (DeKeyser, 2003; Ellis, 2009). However, in the present study, we revealed the effects of implicit statistical learning even at a learning period much earlier than those that had been suggested in previous studies. These results suggest that statistical learning is often implicitly performed in our daily life and that it affects short-term learning and behaviors.

Our results suggest that the N1m peak latencies in the left hemisphere were significantly longer than those in the right hemisphere. In the implicit condition, the participants were provided a silent movie during auditory stimulus presentation, and before the experiments, the participants were presented the questions about the detailed contents of the movie and told that they were supposed to answer those questions following the MEG measurement. These approaches induced the participants to continue paying attention to the story in the movie. In general, right-handed people have left hemispheric dominance for the learning of visual and auditory information (Sakai, 2005), suggesting that our right-handed participants might be left-hemisphere dominant with respect to learning the movie. Because of the dual processing of both the movie and the statistical learning of the tone sequences, the N1m peak latencies to tones in the left hemisphere were significantly longer than those in the right hemisphere.

Few studies have investigated the difference in effects between implicit and explicit learning compared with the difference in effects between implicit and explicit memory (DeKeyser, 2003). One reason for this is that it is more difficult to differentiate between implicitness and explicitness in learning than it is in memory (DeKeyser, 2003; De Jong, 2005; Ellis, 2005, 2009). According to Norris and Ortega (2000), instructional treatment is explicit if rule explanation forms a part of the instruction (deductive instruction) or if learners are asked to continue to focus on the learning (intentional instruction). Conversely, when neither rule presentation nor directions to attend to particular forms are part of a treatment, the treatment is considered implicit. To date, in studies regarding learning, these notions have been pre-eminent: researchers have been divided between explicit and implicit learning on the basis of the forms of instruction (Alanen, 1995; DeKeyser, 2003; Ellis, 1993, 2009; Leow, 1998; Michas & Berry, 1994; Robinson, 1996; Rosa & O'Neill, 1999). Therefore, we have adopted these approaches in this study. First, in the implicit learning condition, participants were shown a silent movie during an auditory stimuli presentation, thus allowing us to assume that they did not pay attention to the tone sequences (Clark and Squire, 1998; Koelsch & Jentschke, 2008; Virtala et al., 2011). In contrast, in the explicit learning condition, a one-second silent period was pseudo-randomly inserted within every thirty successive tones. In this case, the participants were instructed to raise their right hand at every silent period to ensure that they continued to pay attention to the tone sequences. Furthermore, in the implicit learning condition, participants received both intentional and deductive instructions that were entirely unrelated to the learning of the tone sequences. Before the experiments, the questions were disclosed to the participants as deductive instruction, and they were asked to answer the questions after the MEG measurement as intentional instruction. In the explicit learning condition, participants received both intentional and deductive instructions for the learning of the tone sequences. Before the experiments, participants were told that the tones were sequenced on the basis of transitional probabilities as a deductive instruction and that

they would have to take tests on the tone sequences after the experiments as an intentional instruction. These approaches could affect the targeting of learning between the implicit and explicit conditions. Nonetheless, in the MEG analysis, statistical learning effects were detected in both the implicit and explicit learning conditions, although the effects were less pronounced in the implicit condition than in the explicit condition. These findings might suggest that statistical learning is almost independent of the learner's attention. That is, our results support previous studies that suggest that humans can use implicit statistical mechanisms in learning (Ettliger et al., 2011). To date, no conclusion has been reached regarding which form of learning is more efficient. Furthermore, the difference in learning effects does not necessarily indicate superiority of that learning strategy. In other words, in a particular learning task, implicit learning may be more efficient than explicit learning. In our study, the effects of implicit statistical learning were detected at a much earlier stage of learning. This suggests that implicit learning may affect our daily lives. Future research aimed at a comparison between implicit and explicit learning in many learning tasks is necessary to identify the essential roles of implicit and explicit learning. Our results may be important as they revealed differences between implicit and explicit learning conditions and fundamental mechanisms of auditory learning for humans.

The behavioral tasks in the implicit and explicit conditions were different. In the implicit condition, the participants were instructed to focus on the movie, and they were subsequently asked a set of questions about the detailed contents of the movie after the MEG measurement. In the explicit condition, they were instructed to raise their hand at every silent period while paying attention to the tone sequences, and they were asked whether short tone series sounded familiar to them after the MEG measurement. If the familiarity tests were conducted in the implicit as well as the explicit conditions, the experience of the familiarity test in the implicit condition may have functioned as a nonverbal cue and affect learning in the explicit condition. The experiment in the implicit condition was performed prior to the experiment in the explicit condition for every participant. Essentially, behavioral performances with respect to each condition were not comparable in this study.

Temporal resolution in learning effect indexed by the N1m responses is another limitation of this study. In each portion, there were only 30 tones transitioned with lower probability that commonly appeared in every portion of the sequences. On the other hand, at least 30 times of averaging were needed to extract the source-strength waveforms with a sufficient signal-to-noise ratio. We could not trace the temporal profile of this attenuation with higher resolution due to this limitation.

Previous studies revealed that the relative pitch processing of an atonal melody that was not restricted to the diatonic scale was more difficult than the processing of a tonal melody (Dowling & Fujitani, 1971), thus suggesting that the relative pitch processing could be acquired by a domain-specific mechanism for music (i.e., tonality). However, other studies have reported that animals (Wright, 2007), as well as human infants (Trehub & Hannon, 2006), are capable of relative pitch processing. In addition, we confirmed that the relative pitch processing of auditory sequences could occur in implicit learning as well as in explicit learning. These results show that relative pitch processing is available in domain-general statistical learning as well as in domain-specific learning. In short, our results indicate that relative pitch processing is one of the most essential mechanisms innate to humans.

Although Furl et al. (2011) found statistical learning effects on the P2m as well as on the N1m using pure tones, we detected no effects on the P2m. This may result from the general assumption that the P2m resists adaptation more than the N1m, and this tendency is even more pronounced when vowel sounds are used

as the repeated stimuli (Kuriki et al., 2007; Ross & Tremblay, 2009; Teismann et al., 2004). Furthermore, in the study by Ross et al., the adaptation effects on the N1m and P2m responses to repeated stimuli were detected in different time scales and polarity. In that study, 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. Due to this long time constant of P2m modification, we could not detect a difference in the P2m. Thus, the short-term learning effects may be predominant in our results.

Previous studies have also revealed the effects of statistical learning on event-related potentials, using the tasks introduced by Saffran, Johnson, Aslin, and Newport (1999) (Abla et al., 2008; Parakevopoulos et al., 2012). In these studies, the neural responses to tones with higher transitional probability were significantly decreased compared with the responses to tones with lower transitional probability. The tone sequences used in these studies consisted of simple triplets of pure tones or speech sounds, which were used to detect segmentations of words or tones that may be targeted at statistical learning of relatively simple structures. In contrast, our study addressed the statistical learning of higher-order structures with tone sequences based on second-order Markov chains.

In previous studies, MMNs have been used as indices for differential processing of probable and improbable tones (i.e., standards and oddballs) (e.g., Haenschel et al., 2005; Näätänen et al., 2007; Shestakova et al., 2002). There are two types of hypotheses about the neural mechanisms underlying MMN generation. The first hypothesis is that the MMN represents the change detection process involved in a memory-trace effect, which is functionally and spatially distinct from the N1 generation that has predominantly been interpreted as an afferent input population (Haenschel et al., 2005; Näätänen, 1984). The second hypothesis is that the MMN results from the difference in the adaptation of the N1 responses to standard and deviant stimuli (Jääskeläinen et al., 2004). The N1 responses to standard stimuli are gradually attenuated with repetition because of the adaptation of auditory cortical neurons, whereas the N1 responses to deviant stimuli are resistant to adaptation effects. These competing hypotheses indicate that the neural mechanisms underlying MMN generation have not been fully clarified.

In the present study, the N1m responses to tones with lower transitional probability were increased compared with the responses to tones with higher transitional probability. The waveforms in the present study were similar to the MMN from the perspective of the peak amplitudes and latencies. Although the present study aimed to detect learning effects in the neuromagnetic responses to tone sequences with a higher-order structure in which upcoming tones were statistically defined by the history of the tone series, the neural responses to tones with lower transitional probability during statistical learning and the MMN might share neural substrates for deviance detection.

On the other hand, the adaptation effects on amplitudes and latencies were also detected in the neuromagnetic responses in the present study. The N1 responses to the tones with higher transitional probability were significantly attenuated, whereas the N1 responses to the tones with lower transitional probability were resistant to the attenuation. Further studies are necessary to verify the relationship between the MMN and the adaptation effects on the N1 as an index for the differential processing of standard and deviant stimuli. The present study demonstrated that learning statistically transitioned sequences within a higher-order structure as well as oddball sequences was reflected in the N1m responses in both explicit and implicit conditions.

5. Conclusions

The present results suggest that the N1m response could be a marker for statistical learning in both explicit and implicit conditions, although the effects were less pronounced in implicit learning conditions than in explicit learning conditions, and the results indicate that statistically acquired knowledge may persist across spectral shifts in both learning modalities.

Reference

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*, 952–964.
- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In: R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259–302). Honolulu: University of Hawai'i Press.
- Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science*, *280*, 77–81.
- 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: C. J. Doughty, & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford, MA: Blackwell.
- 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. *Learning and Memory*, *12*, 480–490.
- Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, *69*, 338–351.
- Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, *49*, 524–531.
- Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, *27*, 91–113.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141–172.
- Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In: R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philip, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3–25). Bristol, England: Multilingual Matters.
- Ettlinger, M., Margulis, E. H., & Wong, P. C. (2011). Implicit memory in music and language. *Frontiers in Psychology*, *2*, 211. <http://dx.doi.org/10.3389/fpsyg.2011.00211>.
- Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*, 2357–2365.
- 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.
- Gass, S. M., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning*, *53*, 497–545.
- Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelić, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *Journal of Neuroscience*, *25*, 10494–10501.
- Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193–223.
- Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning: Introduction. *Studies in Second Language Acquisition*, *27*, 129–140.
- Jacoby, L. L. (1983). Remembering data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, *22*, 485–508.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levanen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 6809–6814.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, *67*, 971–995.
- Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain Research*, *1212*, 55–62.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kuriki, S., Ohta, K., & Koyama, S. (2007). Persistent responsiveness of long-latency auditory cortical activities in response to repeated stimuli of musical timbre and vowel sounds. *Cerebral Cortex*, *17*, 2725–2732.
- Leow, R. P. (1998). Toward operationalizing the process of attention in second language acquisition: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, *19*, 133–159.
- Markov, A. A. (1971). *Extension of the limit theorems of probability theory to a sum of variables connected in a chain. Dynamic probabilistic systems, volume 1: Markov chains*. New York: John Wiley and Sons (reprinted in Appendix B of: R. Howard).

- Michas, I. C., & Berry, D. C. (1994). Implicit and explicit processes in a second language learning task. *European Journal of Cognitive Psychology*, 6, 357–381.
- Näätänen, R. (1984). In search of a short-duration memory trace of a stimulus in human brain. In: L. Pulkkinen, & P. Lyytinen (Eds.), *Human action and personality. Essays in honour of Martti Takala* (pp. 22–36). Jyväskylä, Finland: University of Jyväskylä.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118, 2544–2590.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14, 481–487.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811–814.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
- 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.
- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition*, 98, 1–11.
- 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.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219–235.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Rosa, E., & O'Neill, M. (1999). Explicitness, intake, and the issue of awareness: Another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511–556.
- Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, 248, 48–59.
- Saffran, J. R. (2003). Absolute pitch in infancy and adulthood: The role of tonal structure. *Developmental Science*, 6, 35–43.
- 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.
- Sakai, K. L. (2005). Language acquisition and brain development. *Science*, 310, 815–819.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501–518.
- Shestakova, A., Brattico, E., Huottilainen, M., Galunov, V., Soloviev, A., Sams, M., et al. (2002). Abstract phoneme representations in the left temporal cortex: Magnetic mismatch negativity study. *Neuroreport*, 13, 1813–1816.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11–26.
- Schmidt, R. (2001). Attention. In: P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
- 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.
- Squire, L. R., Knowlton, B., & Musen, G. (1993). The structure and organization of memory. *Annual Review of Psychology*, 44, 453–495.
- 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.
- Teismann, I. K., Sörös, P., Manemann, E., Ross, B., Pantev, C., & Knecht, S. (2004). Responsiveness to repeated speech stimuli persists in left but not right auditory cortex. *Neuroreport*, 15, 1267–1270.
- Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, 930, 1–16.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100, 73–99.
- Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30, 37–69.
- Virtala, P., Berg, V., Kivioja, M., Purhonen, J., Salmenkivi, M., Paavilainen, P., et al. (2011). The preattentive processing of major vs. minor chords in the human brain: An event-related potential study. *Neuroscience Letters*, 487, 406–410.
- Wright, A. A. (2007). An experimental analysis of memory processing. *Journal of the Experimental Analysis of Behavior*, 88, 405–433.
- Yumoto, M., Matsuda, M., Itoh, K., Uno, A., Karino, S., Saitoh, O., et al. (2005). Auditory imagery mismatch negativity elicited in musicians. *Neuroreport*, 16, 1175–1178.