Perceptual uncertainty modulates auditory statistical learning: A magnetoencephalography study

学位論文

Perceptual uncertainty modulates auditory statistical learning: A magnetoencephalography study

（遷移確率の違いが聴覚統計学習に与える影響：脳磁図を用いた研究）

福島県立医科大学大学院医学研究科

病態神経学分野（脳神経内科学講座）

岡野　智子

Perceptual Uncertainty Modulates Auditory Statistical Learning: A Magnetoencephalography Study

Tomoko Okano[a,b], Tatsuya Daikoku[b,c*], Yoshikazu Ugawa[d], Kazuaki Kanai[a], Masato Yumoto[b,e]

[a] Department of Neurology, Fukushima Medical University, Fukushima, Japan

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

[c] International Research Center for Neurointelligence (WPI-IRCN), The University of Tokyo, Japan

[d] Department of Human Neurophysiology, Fukushima Medical University, Fukushima, Japan

[e] Advanced Medical Science Research Center, Gunma Paz University, Gunma, Japan

*** Correspondence:**

Tatsuya Daikoku: daikoku.tatsuya@mail.u-tokyo.ac.jp

Abstract

Statistical learning allows comprehension of structured information, such as that in language and music. The brain computes a sequence's transition probability and predicts future states to minimise sensory reaction and derive entropy (uncertainty) from sequential information. Neurophysiological studies have revealed that early event-related neural responses (P1 and N1) reflect statistical learning – when the brain encodes transition probability in stimulus sequences, it predicts an upcoming stimulus with a high transition probability and suppresses the early event-related responses to a stimulus with a high transition probability. This amplitude difference between high and low transition probabilities reflects statistical learning effects. However, how a sequence's transition probability ratio affects neural responses contributing to statistical learning effects remains unknown. This study investigated how transition-probability ratios or conditional entropy (uncertainty) in auditory sequences modulate the early event-related neuromagnetic responses of P1m and N1m. Sequence uncertainties were manipulated using three different transition-probability ratios: 90:10%, 80:20%, and 67:33% (conditional entropy: 0.47, 0.72, and 0.92 bits, respectively). Neuromagnetic responses were recorded when participants listened to sequential sounds with these three transition probabilities. Amplitude differences between lower and higher probabilities were larger in sequences with transition-probability ratios of 90:10% and smaller in sequences with those of 67:33%,

compared to sequences with those of 80:20%. This suggests that the transition-probability ratio finely tunes P1m and N1m. Our study also showed larger amplitude differences between frequent- and rare-transition stimuli in P1m than in N1m. This indicates that information about transition-probability differences may be calculated in earlier cognitive processes.

Keywords: entropy; information theory; magnetoencephalography; Markov model; statistical learning; uncertainty

Abbreviations: ANOVA, analysis of variance; ECDs, equivalent current dipoles; ERFs, event-related magnetic fields; ERPs, event-related potentials; MEG, magnetoencephalography; MMN, mismatch negativity; sMMN, statistical MMN

1. Introduction

The human brain extracts statistical regularities from the world around us to learn about the environment. This factor is referred to as statistical learning, which allows us to predict an upcoming phenomenon to minimise prediction errors and resolve perceptual uncertainty (Friston, 2010; Hasson, 2017). Statistical learning of sequences allows people to develop the ability to produce and comprehend structured information, such as in language and music (Saffran et al., 1996). It entails encoding transition probabilities and deriving entropy (i.e., uncertainty) therefrom in sequential information, as in language and music. Statistical learning is an implicit and innate function of the human brain (Perruchet and Pacton, 2006) and is considered an imperative mechanism in neural development (Daikoku, 2019).

Previous studies have revealed that statistical learning is reflected in both neural and behavioural responses (for a review, see Daikoku, 2018). The neural response in statistical learning must be detectable prior to the learner being consciously aware of learning. Notably, electroencephalography and magnetoencephalography (MEG) can directly measure neural activity during statistical learning and can detect it more evidently than observation of behavioural changes (Koelsch et al., 2016; Paraskevopoulos et al., 2012). In the framework of predictive coding (Friston, 2005), when the brain encodes the transition probability distribution of stimulus sequences, it expects

an upcoming stimulus with a high transition probability and suppresses neural activity in response to a predictable stimulus.

The difference in neural activity between predicable (frequent) and unpredictable (rare) stimuli, therefore, may denote early information processing in statistical learning. Researchers regard this amplitude difference as the effectiveness of statistical learning and define it as the statistical learning effect (De Diego Balaguer et al., 2007; Cunillera et al., 2006; Daikoku et al., 2014, 2015; Francois and Schön, 2011; Furl et al., 2011; Sanders et al., 2002). The difference between frequent- and rare-transition stimuli (statistical learning effect) correlates with the degree of transition probability (Koelsch et al., 2016; Pearce and Wiggins, 2012). For example, event-related potentials (ERPs) and event-related magnetic fields (ERFs) have shown significant effects of statistical learning on various neural components from the early latencies (e.g., P50 or P1: Daikoku et al., 2016; Daikoku and Yumoto, 2017; Paraskevopoulos et al., 2012, N100 or N1: Daikoku et al., 2014, 2015; Furl et al., 2011; Sanders et al., 2002) to the late latencies (e.g., P200: Balaguer et al., 2007; Cunillera et al., 2006; Francois and Schön, 2011; Furl et al., 2011, and N400: Balaguer et al., 2007; Cunillera et al., 2006, 2009; François et al., 2013; François and Schön, 2014; Sanders et al., 2002). It is known that different ERP/ERF components have different roles in information processing. For example, the suppression of the ERP/ERF components in early latencies, such as P1m and N1m,

reflects predictions from higher cognitive cortical areas through a top-down pathway (Daikoku, 2018; Skoe et al., 2015). However, despite many ERP/ERF studies on statistical learning, few have investigated how the uncertainty of a sequence (e.g., entropy or transition-probability ratios) modulates the early ERP/ERF components (Koelsch et al., 2016).

In this study, we investigated how the transition-probability ratio or uncertainty of the auditory sequence influences the statistical learning effect on the early ERF components of P1m and N1m, using MEG. The uncertainty of the three sequences was manipulated using three different transition-probability ratios of 90:10%, 80:20%, and 67:33% (conditional entropy: 0.47, 0.72, and 0.92 bits, respectively). In this study, we termed the frequent sequence as high occurrence rate sequence, and the rare sequence, a low occurrence rate sequence. Neuromagnetic responses were recorded as participants listened to the three sequences. We hypothesised that the ERF responses to stimuli with higher transition probabilities would be significantly smaller than those with lower transition probabilities, as demonstrated in several previous studies. Moreover, we assumed that the amplitude difference of neural responses between higher and lower transition probabilities (statistical learning effect) would have an inverse relationship with conditional entropy (uncertainty): the statistical learning effect would be more and less pronounced in the sequences with transition-probability ratios of 90:10% and less pronounced in the sequence with transition-probability ratios of

67:33%, respectively, compared with the sequence with transition-probability ratios of 80:20%. The transition-probability ratio or uncertainty may finely tune the early ERF components of P1m and N1m.

2. Methods

2.1. Participants

Seventeen right-handed (Edinburgh handedness questionnaires; laterality quotient: mean [standard error of mean] = 84.4 [±8.5]), healthy volunteers participated in this study (12 women; age range, 24–45 years). They had neither neurological nor audiological disorders by self-report. Further, they had no formal musical training for longer than 5 years in addition to their regular school lessons. None of the participants possessed absolute pitch. The study was carried out in accordance with the guidelines of the 1964 Declaration of Helsinki and was approved by the Ethics Committee of the University of Tokyo. All participants were well informed about the purpose and safety of this investigation, assured that their data would be protected, and provided written informed consent to participate in this study.

2.2. Stimuli and Experimental Protocol

Three types of 1,500-tone sequences were prepared and are detailed in the S1 Appendix. In each sequence, the 1,500 pure tones that consisted of five pitches in a five-tone equal temperament ($F0 = 350 \times 2^{(n-1)/5}$ Hz, n = 1–5; 350, 402, 462, 531, and 609 Hz) were sequenced with the constraint that the probability of a forthcoming tone was statistically defined by the two most recent tones (second-order Markov models). They were presented with a constant stimulus onset asynchrony of 600 ms (duration, 300 ms; rise/fall, 10/200 ms; binaural presentation, 80 dBSPL intensity). By manipulating the transition-probability ratios of the Markov models, three sequences with different transition-probability ratios of 90:10%, 80:20%, and 67:33% were prepared. Based on information theory, the conditional entropies are 0.47 bits in the sequence with a transition-probability ratio of 90:10%, 0.72 bits in the sequence with a transition-probability ratio of 80:20%, and 0.92 bits in the sequence with a transition-probability ratio of 67:33% (about this formula, see Daikoku, 2018). The three distinct types of Markov chains were used in each of the three sequences because if the same Markov chain was used in all three sequences with different transition-probability ratios, participants may realise the abstract transition pattern even if the transition-probability ratio was distinct. Further, the use of Markov chains was also counterbalanced across participants to avoid the effect of a specific transition pattern of a Markov chain. The three types of diagrams of Markov chains are provided in the S2 Appendix. A representative transition matrix of

the Markov chains is shown in Figure 1. A 600-ms silent period was inserted in a pseudorandom manner within every set of 50 successive tones in the sequence; the participants were instructed to raise their right hand at every silent period in the sequence to confirm that they attended to the tone sequences. We identified the performance of this task by visual inspection and an experimenter recording times of omission. The participants listened to the three types of statistical learning sequences during MEG measurement and took a behavioural test immediately after each statistical learning sequence. In the behavioural tests, they were presented with 10 types of 8-tone series, half of which were sequenced by the same Markov model. They indicated whether each tone series sounded familiar or not. The experiment duration for each participant was about 1 hour.

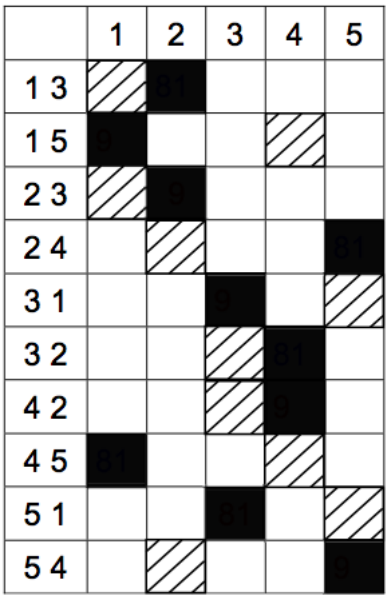

Figure 1. Representative transition matrices of Markov chains used in present study

Rows: the two tones that were most recently presented. Columns: the next tone that might appear. Each pair of two tones (rows) could be followed by one of the five subsequent tones (columns), with a higher (black cells) or a lower (dashed cells) probability. By manipulating the transition-probability ratios of the Markov models, the three sequences with different transition-probability ratios of 90:10%, 80:20%, and 67:33% (conditional entropy: 0.47, 0.72, and 0.92 bits, respectively) were prepared. Three distinct types of Markov chains were used in each of the three sequences because participants might realise the abstract transition pattern if the same Markov chain was used in all sequences. The use of Markov chains was also counterbalanced across participants to avoid the effect of the specific transition pattern of a Markov chain.

2.3. Measurement and analysis

Measurement and analysis were the same as those used in our previous study (Daikoku et al., 2014). Auditory stimuli were sequenced with the STIM2 system (Compumedics Neuroscan, E1 Paso, TX, USA) and were binauralised to the participants' ears at 80 dVSPL through an ER-3A earphone (Etymotic Research, Elk Grove Village, IL, USA). The order of the three sequences was counterbalanced across participants to ensure that specific transition patterns did not interfere with learning in adjacent sequences.

We recorded MEG signals from the participants while they listened to the three sequences. MEG-signal recording and the analyses were performed using the Elekta Neuromag system (Oy, Helsinki, Finland), a 306-channel neuromagnetometer system with 204 planar first-order gradiometers and 102 magnetometers at 102 measuring sites (see S2 Appendix) on a helmet-shaped surface covering the entire scalp with the Elekta Neuromag system (Oy, Helsinki, Finland) in a magnetically shielded room. Epochs induced by the auditory stimuli were filtered online with a 0.1 Hz to 200 Hz range and recorded at a sampling rate of 600 Hz. Epochs with artifacts that exceeded 3 pT/cm or 3 pT for any MEG channel were excluded from the analyses. Contamination from environmental noise was reduced using the temporally extended signal space separation method with a buffer length of 10 s and a correlation limit of 0.980. All responses to tones in the three sequences

were averaged for each participant to evaluate the reliability of the single equivalent current dipoles (ECDs) for each of the P1m and N1m components. The averaged responses were filtered offline with a 2–40 Hz band-pass. The baseline for magnetic signals in each MEG channel was defined by the mean amplitude in the pre-stimulus period from −100 to 0 ms. The ECDs for P1m and N1m responses in each hemisphere were separately estimated at the peak latency using 44 temporal channels without magnetometers in each participant (S2 Appendix). Participants who demonstrated poor ECD estimation, with a goodness-of-fit below 70% in either the left or right hemisphere, were not used in further analyses. Consequently, the P1m and N1m components were studied in 15 and 12 participants, respectively (see S2 Appendix).

Selective response averaging was performed separately for each phase, first (1-500 tones), middle (501-1,000 tones), and last (1,001-1,500 tones), in 1,500-tone sequences, and were further divided depending on higher and lower transition probabilities. To illustrate this process, in a sequence with transition-probability ratios of 90:10%, 450 trials of frequent-transition tones (i.e., tones with 90% transition probability) and 50 trials of rare-transition tones were averaged in each phase. In a sequence with transition-probability ratios of 80:20%, 400 trials of frequent-transition and 100 trials of rare-transition tones were averaged in each phase. In the sequence with transition-probability ratios of 67:33%, 335 trials of frequent-transition and 165 trials of rare-transition tones

were averaged in each phase. The analysis window was defined as 0−500 ms. Using the ECDs, the source strength of P1m and N1m in each hemisphere was calculated based on the averaged selective response. Then, the peak amplitude and latency were measured for each condition in each subject. We performed a 2 (hemisphere: right and left) × 3 (phase: first, middle, and last) × 2 (occurrence rate: high and low) repeated-measures analysis of variance (ANOVA) with peak amplitude and latency of source strength of each component and sequence. Furthermore, we performed an ANOVA with the logit values of the familiarity ratios in the behavioural test. Significance levels were set at $p = 0.05$. In post hoc analyses, we used a t-test with Bonferroni correction to compensate for multiple comparisons. All statistical analyses were performed using the jamovi statistical platform (Version 1.2; The jamovi project, 2020; https://www.jamovi.org).

3. Results

3.1. MEG

We evaluated how the sequence transition-probability ratios influenced P1m and N1m responses. Stimuli with lower transition probabilities elicited stronger P1m and N1m responses than those with higher transition probabilities (Figures 2 and 3, S4 Appendix). This amplitude difference (higher vs. lower transition probabilities) was significant, for both the N1m and P1m components, in

the sequence with a transition-probability ratio of 90:10%, whereas no significant difference was observed in the sequence with a transition-probability ratio of 67:33%. During exposure to the sequence with a transition-probability ratio of 80:20%, an amplitude difference was detected in P1m but not N1m (see S3 Appendix for the mean P1m and N1m amplitudes and latencies). These observations were confirmed by ANOVAs.

3.1.1. P1m

The grand-averaged source-strength waveforms of the P1m responses in the left and right hemispheres are shown in Figure 2. The response amplitude difference between frequent- and rare-transition stimuli (i.e., stimuli with high and low transition probabilities) was observed in the sequences with transition-probability ratios of 90:10% (Figure 2b) and 80:20% (Figure 2c) but not 67:33% (Figure 2d). This impression was confirmed by statistical analysis.

An ANOVA of the peak P1m amplitude revealed a significant main effect of the occurrence rate in the sequences with 90:10% ($F[1,14] = 4.714$, $p = 0.048$) and 80:20% ($F[1, 14] = 5.718$, $p = 0.031$) but not the 67:33% sequence. The P1m responses to standard- or frequent-transition stimuli were significantly smaller than those to rare-transition stimuli. This finding is consistent with those of several previous ERP and ERF studies.

The factor of hemisphere significantly affected the P1m peak amplitude (90:10%: $F[1, 14] = 8.367$, $p = 0.012$; 80:20%: $F[1, 14] = 6.740$, $p = 0.021$; 67:33%: $F[1, 14] = 8.367$, $p = 0.012$). The peak amplitudes of P1m were significantly greater in the left hemisphere than in the right hemisphere. The main phase effect on the P1m peak amplitudes was significant in the 90:10% and 67:33% sequences ($F[2, 28] = 5.081$, $p = 0.013$). The post hoc test on the phase revealed that the P1m peak amplitudes for tones in the initial phase were significantly smaller than those in the last phase (90:10%: $p = 0.024$, 67:33%: $p = 0.035$). Furthermore, the hemisphere-phase interaction of the P1m peak amplitudes was significant in the 67:33% sequence ($F[2, 28] = 4.100$, $p = 0.027$). The post hoc test on the hemisphere-phase interaction revealed two significant results. Firstly, the P1m peak amplitudes in all phases were significantly greater in the left hemisphere than in the right hemisphere (initial: $p = 0.019$, middle: $p = 0.046$, final: $p = 0.006$). Secondly, in the left hemisphere, the P1m peak amplitudes in the initial phase were significantly smaller than those in the final phase.

The main phase effect on the P1m peak latencies was significant in the 90:10% sequence ($F[2, 28] = 7.135$, $p = 0.003$) and 80:20% sequence ($F[2, 28] = 4.448$, $p = 0.021$). The post hoc test showed that the P1m peak latencies in the initial phase were significantly shorter than those in the middle and final phases in the 90:10% sequence (middle: $p = 0.018$, final phase: $p = 0.004$).

Moreover, in the case of the 80:20% sequence, the P1m peak latencies in the initial phase were shorter than those in the final phase ($p = 0.031$).

The time courses of the differences in the peak amplitudes of P1m responses (i.e., the amplitude of a response to a rare-transition stimulus compared to that of a frequent-transition stimulus) in the left and right hemispheres are summarised in the S4 Appendix. The ANOVAs showed no significant differences in the statistical learning effects by phase. No other significant differences were detected in the ANOVAs for P1m amplitude or latency.

3.1.2. N1m

The grand-averaged source-strength waveforms of the N1m responses in the left and right hemispheres are shown in the S4 Appendix. In the sequence with the transition-probability ratio of 90:10%, the N1m peak amplitudes for the high occurrence rate tones were smaller than those for the low occurrence rate tones in the final phase ($p = 0.040$). The main phase effect on the N1m peak amplitude was significant (90:10%: $F[2, 22] = 18.773$, $p = 0.000$, 80:20%: $F[2, 22] = 5.052$, $p = 0.016$, 67:33%: $F[2, 22] = 11.780$, $p < 0.001$). The post hoc test revealed that the N1m peak amplitudes for tones in the initial phase were significantly larger than those in the middle phase

(90:10%, $p < 0.001$; 67:33%, $p = 0.001$) and in the final phase (90:10%: $p = 0.002$, 67:33%: $p = 0.011$).

The N1m peak latency was significantly affected by the occurrence rate in the 90:10% ($F[1, 11] = 15.994$, $p = 0.002$) and 80:20% sequences ($F[1, 11] = 6.221$, $p = 0.030$) but not in the 67:33% sequence. The N1m peak latencies for frequent-transition tones were significantly earlier than those for rare-transition tones. The main phase effect on the N1m peak latency was significant in the 90:10% ($F[2, 22] = 6.784$, $p = 0.000$) and 67:33% sequences ($F[2, 22] = 10.355$, $p = 0.001$). The post hoc test revealed that the N1m peak latencies in the initial phase were significantly shorter than those in the final phase in the 90:10% sequence ($p = 0.024$). Moreover, in the case of the 67:33% sequence, the N1m peak latencies in the initial phase were shorter than those in the middle and final phases ($p = 0.005$ and 0.018, respectively).

The time courses of the statistical learning effects on the N1m peak (i.e., the amplitude of the response to a rare-transition stimulus compared to that of a frequent-transition stimulus) are shown in Figure 3. The abscissa shows the learning phase, and the ordinate shows the statistical learning effect. The upper row shows the sequence with a transition-probability ratio of 90:10% (Figure 3a), the middle row 80:20% (Figure 3b), and the bottom row 67:33% (Figure 3c). The statistical learning effect became larger in the later phases in all figures. An ANOVA showed a

significant phase-transition probability interaction ($F[2, 22] = 7.538$, $p = 0.003$). The phase significantly affected the statistical learning effect only in the transition probability of the 90:10% sequence. The post hoc test revealed that the statistical learning effect was significantly larger in the last phase than in the initial and middle phases in the sequence with a transition probability of 90:10%. No other significant differences were detected by the ANOVAs for either N1m amplitude or latency.

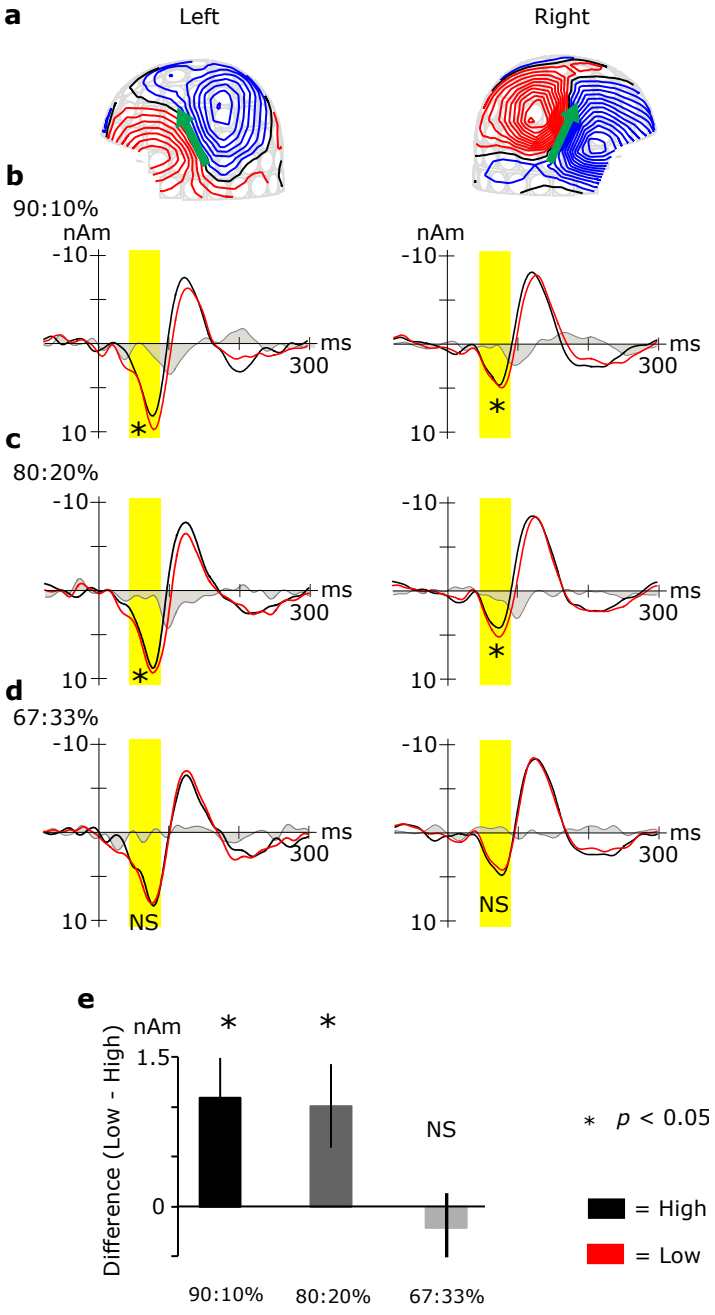
a
Left
Right
b
90:10%
nAm
-10
10
ms
300
*
c
80:20%
d
67:33%
NS
e
Difference (Low - High)
1.5
0
NS
90:10%
80:20%
67:33%
* p < 0.05
= High
= Low

Figure 2. P1m source

ISO filed contour maps for P1m source **(a)** are shown for each hemisphere (top). Outflux (red lines) and influx (blue lines) are stepped by 20 ft. The green arrows represent ECDs. The N1m responses were separately modelled as single attention ECDs in each hemisphere. Grand-averaged waveforms for the P1m responses to frequent- (black) and rare-transition (red) stimuli in the sequences with transition-probability ratios of 90:10% **(b)**, 80:20% **(c)**, and 67:33% **(d)**. The ordinate shows the P1m amplitude difference between the frequent- and rare-transition stimuli (statistical learning effect) **(e)**.

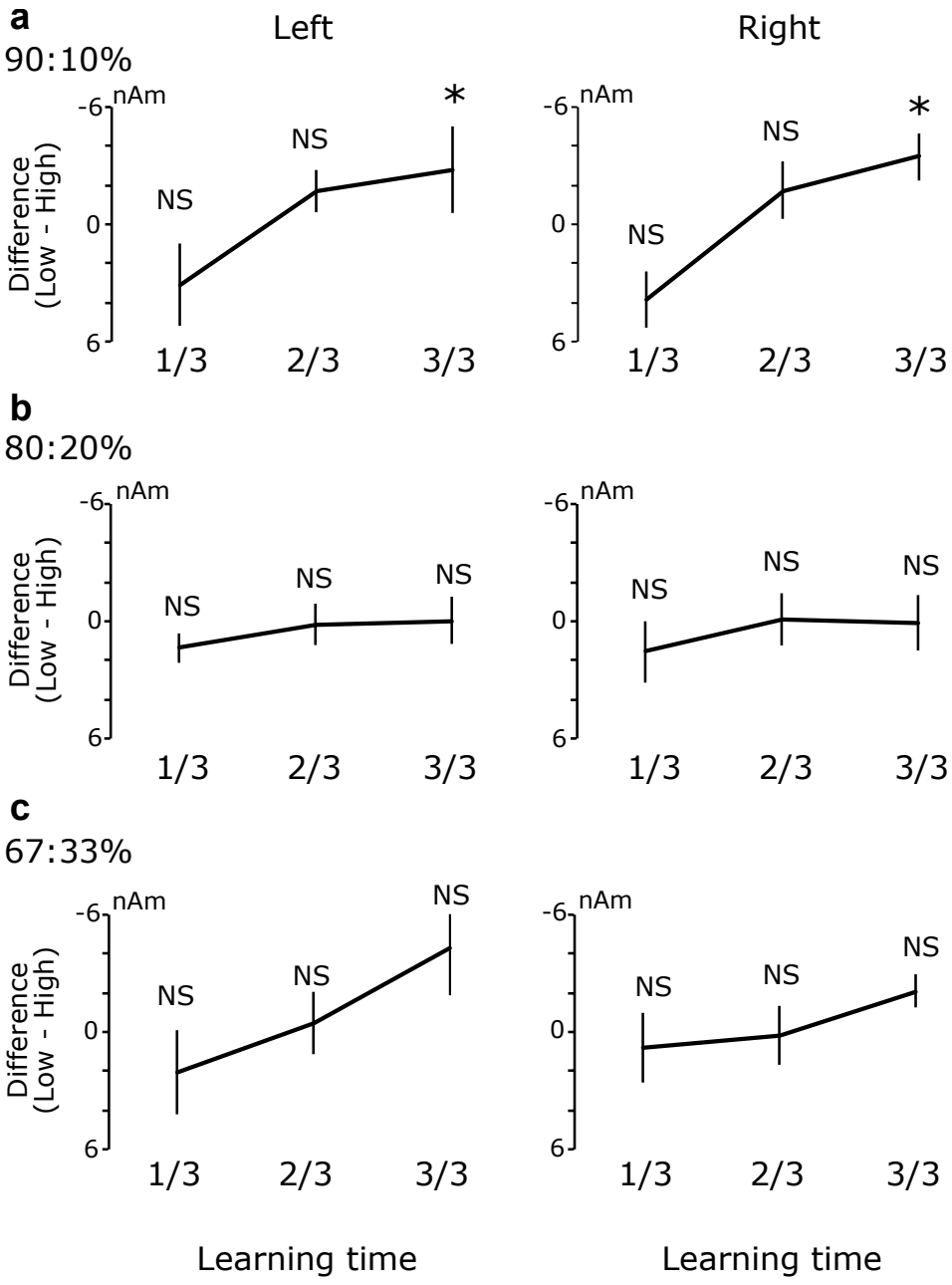
a
Left
Right
90:10%
nAm
-6
0
6
Difference
(Low - High)
NS
NS
*
1/3
2/3
3/3
b
80:20%
c
67:33%
Learning time

Figure 3. Differences in N1m amplitudes between frequent- and rare-transition stimuli (statistical learning effect)

The ordinate shows the N1m peak amplitude difference between the rare- and frequent-transition stimuli (i.e., response to a rare-transition stimulus compared to that of a frequent-transition stimulus). The abscissa shows the phase of the stimulus. The results of the 90:10% sequence is shown in the top row, the 80:20% sequence in the middle row, and the 67:33% sequence in the bottom row. The left columns are responses over the left hemisphere, and the right columns are those over the right hemisphere.

3.2. Behavioural results

All participants correctly raised their right hand at every silent period, further confirming that they had attended to the tone sequences. The correct answer rates of the familiarity tests were significantly above the chance level for all three sequences ($p < 0.001$). They were converted into empirical logit values, which were used for the ANOVA. The ANOVA showed no significant effect of the transition-probability ratio on the correct answer rate (i.e., conditional entropies).

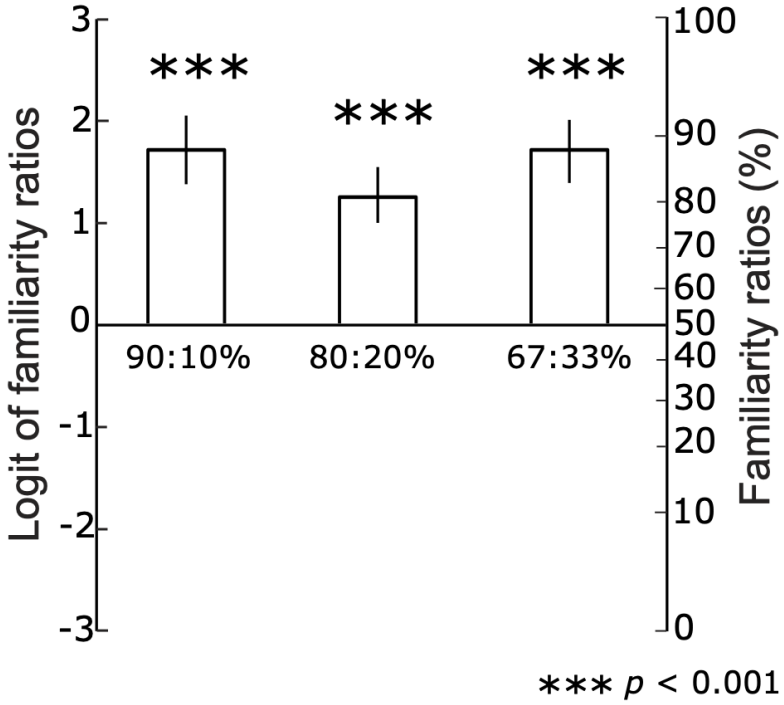

Figure 4. Results of familiarity tests to assess statistical learning

The participants answered correctly in all three sequences (i.e., conditional entropies and uncertainty).

4. Discussion

We studied how transition-probability ratio or uncertainty in the auditory stimulus sequence (conditional entropy) modulates the statistical learning effects of early ERF components, P1m and N1m. MEG was recorded when participants were presented with three sound sequences with different transition-probability ratios (90:10%, 80:20%, and 67:33%; conditional entropy: 0.47, 0.72, and 0.92 bits, respectively). We showed that stimuli with lower transition probabilities elicited stronger neural P1m and N1m responses than those with higher transition probabilities (Figures 2 and 3). This finding is consistent with several previous ERP and ERF studies (e.g., Abla et al., 2008; Daikoku and Yumoto, 2019; Daikoku, 2018; François et al., 2013; Tsogli et al., 2019).

In addition, the amplitude difference between frequent- and rare-transition stimuli (statistical learning effect) was most pronounced in the sequence with transition-probability ratios of 90:10% and lower conditional entropy (0.47). This indicates that the transition probability ratio finely tunes each neural component. Exposure to the sequence with a transition-probability ratio of 90:10% elicited a significant learning effect on both P1m and N1m responses, whereas no significant statistical learning effect was detected during exposure to the sequence with a transition-probability

ratio of 67:33%. Exposure to the 80:20% sequence elicited a statistical learning effect on P1m but not N1m.

We suggest the following plausible explanation for these results. In the 90:10% sequence, the brain is surprised by rare-transition stimuli because the subjects mostly expect frequent-stimuli. Then, the difference between the rare- and frequent-transition stimuli should be large. In the 67:33% sequence, however, surprising effects should be small because rare- and frequent-transition stimuli appear at similar probabilities. Then, the significant difference in the response amplitude between the rare- and frequent-transition stimuli should not be present.

Our finding of the phase effect on the statistical learning is also compatible with this speculation. The statistical learning effect became larger in the late phase than in the earlier phases only in sequences with transition-probability ratios of 90:10% (Figure 3). The learning process should enhance the fine-tuning more in the 90:10% task, which makes N1 and P1 responses larger in the late phase of learning. On the contrary, in sequences with transition-probability ratios of 67:33%, the learning process must not be effective because rare- and frequent-transition stimuli appear at similar probabilities. The statistical learning effect (amplitude difference), therefore, may not become considerably larger in the late phase of learning with transition-probability ratios of 67:33%.

Previous studies have compared the statistical learning effects on P1m and N1m (Paraskevopoulos et al., 2012; Daikoku et al., 2020) and have speculated about the related neural networks (Evangelos Paraskevopoulos, Chalas, Kartsidis, Wollbrink, & Bamidis, 2018). They have suggested that the P1m may reflect more proficiency or individual statistical learning ability compared to N1m. Further, the connectivity analysis showed that those participants with P1m statistical learning effects had enhanced neural network efficiency in the superior-temporal and inferior-frontal gyri. Thus, it is suggested that the P1 component may represent processing efficiency. The present result of a large sequence learning effect from strong contrast learning is compatible with the above hypothesis on the processing efficiency.

According to previous studies, the complexity or uncertainty of sequences influences the performance of statistical learning, particularly in people with developmental learning disorders such as dyslexia (Du and Kelly, 2013) and amusia (Omigie et al., 2012, 2013; Omigie and Stewart, 2011). This implies that people with developmental learning disorders have impaired statistical learning of complex but not simple tasks. Our results could not reflect the corresponding modulation of the neural responses congruent with the probability ratios used. Thus, our findings may suggest that the neurophysiological observation of ERP/ERF is a more sensitive method for evaluating the performance and proficiency of statistical learning than observation of behavioural effects (Koelsch,

Busch, Jentschke, & Rohrmeier, 2016; Evangelos Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). Our neural findings may offer novel neurophysiological insights into the mechanisms underlying statistical learning – the P1m components underlying statistical learning may be used to detect hidden statistical learning impairments and their associated developmental abnormalities.

Previous studies found that the mismatch negativity (MMN) observed 100–200 ms after stimulus onset is also sensitive to the difference in the transition probability (François et al., 2017; Koelsch et al., 2016; Moldwin et al., 2017). Koelsch et al. (2016) found similar results of MMN to our findings on P1 and N1. They found that the MMN amplitudes showed an inverse relation with transition probabilities: a stimulus with lower transition probability elicits higher MMN amplitude. Tsogli et al. (2019) also found that statistical deviants elicited prominent mismatch ERP responses ~150–250 ms after stimulus onset. The authors referred to the ERP effects due to statistical deviants as statistical MMN (sMMN), a term that distinguishes such effects from the traditional MMN elicited by a deviant based on appearance probability instead of transition probability (Paavilainen et al., 1989; Sams et al., 1985). Our recent study found that a positive component before MMN or N1 (around 120 ms after stimulus onset) reflected more processing efficiency (Daikoku et al., 2021). Thus, the present and previous results are both compatible with the hypothesis that the early positive components underlying statistical learning could reflect processing efficiency.

From the viewpoint of the information theory (Shannon, 1948), the perceptual uncertainty of each sequence can be evaluated by conditional entropy. Conditional entropy can be calculated from the transition-probability distribution, interpreted as the average number of surprising or uncertain outcomes. The degree of conditional entropy modulates the predictability of each stimulus of sequences in statistical learning (Agres et al., 2018; Hasson, 2017; Nastase et al., 2014). Several studies have indicated that a sequence with higher uncertainty can be interpreted as difficult to learn (Daikoku, 2018; Hansen and Pearce, 2014; Harrison et al., 2006). Our findings showed neurophysiological evidence that perceptual uncertainty manipulated by transition probability influences statistical learning effects. The amplitude difference of neural responses between higher and lower transition probabilities will have an inverse relation with conditional entropy (uncertainty). However, the transition-probability ratio is not the only indicator of difficulty. For example, the number of stimuli could also affect the conditional entropy even if the transition-probability ratio were the same. Furthermore, types of stimuli (e.g., timbre, length, and loudness) should also influence the difficulty in statistical learning. Further research is needed to investigate how statistical learning effects are modulated by different methods of uncertainty or difficulty manipulation. Nevertheless, our study is the first to reveal that the statistical learning effect on ERF correlates negatively with perceptual uncertainty.

We also showed the dependence of the ERP effects of sequence learning on transition probability; strong contrast sequences have stronger effects on ERPs. These findings suggest that early information processing may occur in an easy task with low entropy and may not be effective in difficult tasks with high entropy. Boutros et al. (1995) demonstrated that P1 responses to frequent-transition stimuli decreased during auditory oddball paradigms with a probability ratio of 90% vs. 10% and 80% vs. 20% (Boutros and Belger, 1999). They concluded that each stimulus generates local inhibitory activity that suppresses the neural response to a second identical stimulus to prepare rare stimuli. In contrast, the novel stimulus triggers the disinhibition of the above pre-set inhibitory circuit and allows the neurons to respond effectively to the deviant stimulus. This gating function should be detected by the amplitude difference in the earlier component of P1. The above gating must readily occur in easy tasks and occur infrequently in difficult tasks.

In conclusion, this study revealed how the transition-probability ratio or the conditional entropy (uncertainty) in the auditory sequence modulates statistical learning effects on the early ERFs of P1m and N1m. A strong contrast sequence evoked a large learning effect on ERPs (P1m, N1m). It suggests that the transition-probability ratio or uncertainty must be calculated at the early timings of P1m or N1m and that the transition-probability ratio or the uncertainty may finely tune P1m and N1m responses. We emphasise that novel experimental paradigms of statistical learning

applied in this study may shed new light on understanding how our brain acquires sequential knowledge with different uncertainties and complexities, such as that found in language and music.

Acknowledgments

This work was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (KAKENHI) [Grant Number 20K22676] and the Suntory Foundation. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

The experimental paradigms of the present study were considered by MY and TD. TD created the paradigms, and TO recruited participants and collected the data. MY proposed methodologies for MEG data analyses. TD and TO analysed both MEG and behavioural data. All authors discussed how the results could neurophysiologically and psychologically be interpreted, and reviewed and revised the manuscript. TD, TO, and YU prepared the figures and wrote the main manuscript text.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

References

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

Agres, K., Abdallah, S., Pearce, M., 2018. Information-theoretic properties of auditory sequences dynamically influence expectation and memory. Cogn. Sci. 42, 43–76. https://doi.org/10.1111/cogs.12477.

Boutros, N.N., Belger, A., 1999. Midlatency evoked potentials attenuation and augmentation reflect different aspects of sensory gating. Biol. Psychiatry 45, 917–922. https://doi.org/10.1016/S0006-3223(98)00253-4.

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

Cunillera, T., Toro, J.M., Sebastián-gallés, N., 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. https://doi.org/10.1016/j.brainres.2006.09.046.

Daikoku, T., 2018. Neurophysiological markers of statistical learning in music and language: Hierarchy, entropy, and uncertainty. Brain Sci 8, 114. https://doi.org/10.3390/brainsci8060114.

Daikoku, T., 2019. Implicit learning in the developing brain: An exploration of ERP indices for developmental disorders. Clin Neurophysiol 130, 2166–2168. https://doi.org/10.1016/j.clinph.2019.09.001.

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

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

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

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

Daikoku T, Yumoto M. 2020. Musical Expertise Facilitates Statistical Learning of Rhythm and the Perceptive Uncertainty: A cross-cultural Study. Neuropsychologia 146, 107553. https://doi.org/10.1016/j.neuropsychologia.2020.107553

Daikoku T, Jentschke S, Tsogl B, Bergström K, Lachmann T, Ahissa M, Koelsch S., 2021. Statistical learning in the developing brain. Neuroscience and music Aarhus, Denmark.

R, Toro JM, Rodriguez-Fornells A, Bachoud-Lévi AC., 2007. Different neurophysiological mechanisms underlying word and rule extraction from speech. PLoS One 2, e1175. https://doi.org/10.1371/journal.pone.0001175.

Du, W., Kelly, S.W., 2013. Implicit sequence learning in dyslexia: A within-sequence comparison of first- and higher-order information. Ann Dyslexia 63, 154–170. https://doi.org/10.1007/s11881-012-0077-1.

François, C., Chobert, J., Besson, M., Schön, D., 2013. Music training for the development of speech segmentation. Cereb Cortex 23, 2038–2043. https://doi.org/10.1093/cercor/bhs180 23(9).

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

François, C., Schön, D., 2014. Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. Hear Res 308, 122–128. https://doi.org/10.1016/j.heares.2013.08.018.

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

Friston, K., 2005. A theory of cortical responses. Philos Trans R Soc Lond B 360, 815–836. https://doi.org/10.1098/rstb.2005.1622.

Friston, K., 2010. The free-energy principle: A unified brain theory? Nat Rev Neurosci 11, 127–138. https://doi.org/10.1038/nrn2787.

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. https://doi.org/10.1016/j.neuroimage.2010.10.038.

Hansen, N.C., Pearce, M.T., 2014. Predictive uncertainty in auditory sequence processing. Front Psychol 5, 1052. https://doi.org/10.3389/fpsyg.2014.01052.

Harrison, L.M., Duggins, A., Friston, K.J., 2006. Encoding uncertainty in the hippocampus. Neural Netw 19, 535–546. https://doi.org/10.1016/j.neunet.2005.11.002.

Hasson, U., 2017. The neurobiology of uncertainty: Implications for statistical learning. Philos Trans R Soc Lond B Biol Sci 372, 20160048. https://doi.org/ 10.1098/rstb.2016.0048.

Koelsch, S., Busch, T., Jentschke, S., Rohrmeier, M., 2016. Under the hood of statistical learning: A statistical MMN reflects the magnitude of transition probabilities in auditory sequences. Sci Rep 6, 19741. https://doi.org/10.1038/srep19741.

Moldwin T, Schwartz O, Sussman ES., 2017. Statistical learning of melodic patterns influences the brain' s response to wrong notes. J Cog Neurosci 29, 2114–2122. https://doi.org/10.1162/jocn_a_01181.

Nastase, S., Iacovella, V., Hasson, U., 2014. Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. Hum Brain Mapp 35, 1111–1128. https://doi.org/10.1002/hbm.22238.

Omigie, D., Pearce, M.T., Stewart, L., 2012. Tracking of pitch probabilities in congenital amusia. Neuropsychologia 50, 1483–1493. https://doi.org/10.1016/j.neuropsychologia.2012.02.034.

Omigie, D., Pearce, M.T., Williamson, V.J., Stewart, L., 2013. Electrophysiological correlates of melodic processing in congenital amusia. Neuropsychologia 51, 1749–1762. https://doi.org/10.1016/j.neuropsychologia.2013.05.010.

Omigie, D., Stewart, L., 2011. Preserved statistical learning of tonal and linguistic material in congenital amusia. Front Psychol 2, 109. https://doi.org/10.3389/fpsyg.2011.00109.

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

Pearce, M.T., Wiggins, G.A., 2012. Auditory expectation: The information dynamics of music perception and cognition. Top Cogn Sci 4, 625–652. https://doi.org/10.1111/j.1756-8765.2012.01214.x.

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

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

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

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst Tech J 27, 623–656. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Skoe, E., Krizman, J., Spitzer, E., & Kraus, N., 2015. Prior Experience Biases Subcortical Sensitivity to Sound Patterns. Journal of Cognitive Neuroscience, 27, 124–140. https://doi.org/10.1162/jocn_a_00691

Tsogli, V., Jentschke, S., Daikoku, T., Koelsch, S., 2019. When the statistical MMN meets the physical MMN. Sci Rep 9, 5563. https://doi.org/10.1038/s41598-019-42066-4.