35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Neuromagnetic Auditory Steady State Response to Chords: Effect of Frequency Ratio*

Asuka Otsuka, Masato Yumoto, Shinya Kuriki and Seiji Nakagawa

Abstract— **Perceptual degree of consonance or dissonance of a chord is known to be varied as a function of frequency ratio between tones composing the chord. It has been indicated that generation of a sense of dissonance is associated with the auditory steady-state response (ASSR) phase-locked to difference frequencies which are salient in the chords with complex frequency ratios. This study further investigated how the neuromagnetic ASSR would be modulated as a function of the frequency ratio when the acoustic properties of the difference frequency, to which the ASSR was synchronized, was identical in terms of its number, energy and frequency. Neuronal frequency characteristics intrinsic to the ASSR were compensated by utilizing responses to a SAM (Sinusoidally Amplitude Modulated) chirp tone sweeping through the corresponding frequency range. The results showed that ASSR was significantly smaller for the chords with simple frequency ratios than for those with complex frequency ratios. It indicates that the basic neuronal correlates underlying the sensation of consonance/dissonance might be associated with the attenuation rate applied to encode the input information through the afferent auditory pathway. Attentional gating of the thalamo-cortical function might also be one of the factors.**

Keywords: magnetoencephalography, auditory cortex, frequency characteristics, consonance, dissonance, human.

I. Introduction

Perceptual degree of consonance or dissonance is varied as a function of frequency ratio between the tones composing the chord. Chords which are capable of simple integral frequency ratio such as 2:1 (octave), 2:3 (perfect 5th), 3:4 (perfect 4th) and 4:5 (major 3rd) are perceived consonant whereas chords capable of complex frequency ratio such as 15:16 (minor 2nd) are perceived dissonant. Perception of dissonance is attributed to sensation of beats or roughness, which is induced at difference frequency (Δf) due to interference between adjacent frequencies processed within 1/4 width of the auditory critical band [1-3].

*Resrach supported by the Funding Program for Next-Generation World-Leading Researchers provided by the Cabinet Office, Government of Japan.

A. Otsuka is with Health Research Institute, the National Institute of Advanced Industrial Science and Technology (AIST), Osaka 563-8577 Japan (phone: +81-72-751-8297, fax:+81-72-751-9517, e-mail: asuka.ootsuka@aist.go.jp).

M. Yumoto is with the Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033 Japan (e-mail: yumoto-tky@umin.ac.jp).

S. Kuriki is with Research Center for Advanced Technologies, Tokyo Denki University, Chiba 270-1382 Japan (e-mail: skuriki@rcat.dendai.ac.jp).

S. Nakagawa is with Health Research Institute, The National Institute of Advanced Industrial Science and Technology (AIST), Osaka 563-8577 Japan (e-mail: s-nakagawa@aist.go.jp).

The neuronal processing of Δf is reflected in the auditory cortical steady-state response (ASSR), an oscillatory evoked response phase-locked to the temporal envelope periodically amplitude-modulated at rates of Δf [4-7]. Intracranial recording revealed that the ASSR was correlated to the energy of Δf and exhibited larger activities for dissonant than consonant chords [8]. However, it was unclear whether the results reflected the cognitive function or the total amount of the Δf contained in the stimulus chords because the acoustic features of the stimulus chords differed in terms of carrier frequency (f_c), amplitude modulation rate (f_m), i.e., Δf, number of Δf and their spectral energy level. The ASSR is known to be varied as functions of sound intensity level [9-11], f_c [11-13] and f_m [10-11,14]. It is therefore considered necessary to manipulate the acoustic features of the stimulus sounds as well as to compensate the neuronal characteristics intrinsic to the ASSR when examining the cognitive function as indexed by the ASSR.

The present study aimed at exploring how the strength of the neuromagnetic ASSR would be modulated as a function of frequency ratio of three-note chords. The carrier frequencies of the chords were determined to have the same Δf in terms of the frequency and energy level. The sound pressure levels of the stimulus chords were calibrated to be equal nearby the ear drums of each subject [13]. The frequency characteristics of the ASSR were compensated by utilizing the response to a sinusoidally amplitude-modulated (SAM) chirp tone sweeping through the corresponding frequency range [13].

II. Methods

A. Stimulus chords

Seven kinds of stimulus chords composed of three sinusoids were prepared in duration of 3.75 seconds. The frequency ratio was varied systematically from 2:3:4, 4:5:6, 6:7:8, 8:9:10, 10:11:12, 12:13:14 through to 14:15:16. They could be also characterized by the interval width of 3rd, 5th, 7th and 9th through to 15th note in the harmonic overtones. The carrier frequencies of each chord, ranging from 128 to 624.51 Hz, were determined for the envelope to be amplitude modulated equally at Δf (f_2-f_1 and f_3-f_2) of 41.6 Hz. A sinusoidally amplitude modulated (SAM) chirp tone exponentially ascended or descended between 128 and 624.51 Hz in duration of 3.75 seconds was also prepared as a control to capture the carrier frequency characteristics of the ASSR. All stimulus tones were synthesized digitally on PC (Precision M2400, DELL Inc., Round Rock, TX, USA) at a sampling rate of 96 kHz with 32 bit resolution (Matlab, The MathWorks, Inc., Natick, MA, USA).

B. Stimulus delivery and intensity calibration

Stimulus delivery for the subjects was prepared in the magnetically shielded room for the left and right ears, respectively. First, a soft silicone rubber probe tube (0.95 mm OD x 0.58 mm ID x 76 mm long (approx. 3.0")) (ER7-14C, Etymotic Research. Inc., Elk Grove Village, IL, USA) was inserted into the ear canals for a depth of about 30 mm from the mastoid tip to a position where the subjects claimed a rustling noise, and fixated with surgical tapes. The outer end of the probe tube was attached to the microphone (ER7c, Etymotic Research. Inc., Elk Grove Village, IL, USA). Next, 13 mm foam eartips (ER1-14A, Etymotic Research. Inc., Elk Grove Village, IL, USA) attached to the tip of one meter plastic tube were inserted into the ear canals for a depth of about 20 mm. Then, the lateral end of the ear canals, together with the tubes and eartips, were sealed tightly with soft silicon ear plugs (Insta-putty, Insta-Mold Products, Inc., Oak, PA, USA).

White noise and a chord composed of 31 sinusoidal frequency components were delivered to and recorded from each ear using the probe microphone as displayed in Figure 1. The recorded signals were fast-Fourier transformed at real time (Audition 1.0, Adobe Systems Inc., San Jose, CA, USA). The spectral characteristics of the white noise were calibrated to be broadly flat by using low-pass, high-pass and band-pass filters. Then, the peak intensities of the spectral power at the 31 frequencies were fine-tuned manually using 31-band graphic equalizer function (DEQ2496, Behringer GmbH, Kirchardt, Germany), resulted in delivering all the frequency components of all the chords equally at 70 dB SPL with deviation of +/- 0.5 dB SPL. The transmission delay of 7.83 ms was identified by the time point of zero crossing of the Hilbert-transformed positive and negative envelopes and compensated by a shift of the trigger signal. The ER7c microphone was detached and removed out of the magnetically shielded room after the calibration.

C. Perceptual evaluation on degree of consonance

Before the MEG measurement, perceptual degree of consonance was evaluated for each of seven stimulus chords on a two-point scale. The stimulus chords were presented to the subjects in pseudo random order at the inter-stimulus interval (ISI) of 1.25 seconds and repeated for three sessions. The presentation order of the chords was randomized per subject per session.

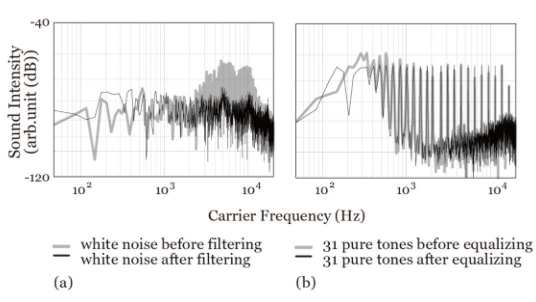

Figure 1. Sound intensity level calibration using (a) white noise and (b) a chord composed of 31 frequency components. Signals recorded in the ear canal using a probe microphone before and after the calibration are overlaid in gray and black lines, respectively.

D. MEG Signal Measurement

Five right-handed subjects with a mean age of 36.8 (SD=7.42) years, normal hearing and no history of otological or neurological disorder participated in the experiment. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of The National Institute of Advanced Industrial Science and Technology (AIST), Japan. The informed consent was obtained from each subject after the purpose and procedures of the experiment were fully explained.

The MEG signals were recorded in a magnetically shielded room using a whole-head SQUID gradiometer (Neuromag-122, Elekta Neuromag Oy, Helsinki, Finland) that detects a radial component of the magnetic field at 61 measurement locations and yields 122-channel gradient signals (ΔBz/Dx, ΔBz/Dy) in two orthogonal directions on the surface over the skull. The signals were recorded at 0.03–200 Hz of an analog frequency band with a sampling frequency of 618 Hz with 24 bit resolution. The stimulus chords were presented repetitively for 150 times at the ISI (Inter-Stimulus Interval) of 1.25 seconds. The ascending and descending SAM chirp tones were presented 75 times each to acquire the signals for 150 epochs in total. During the recording, the subjects were requested to watch a silent film with subtitles at their own choice and ignore the stimulus tones in order to reduce the effect of attentional modulations.

E. Signal processing

The recorded MEG signal was spatiotemporally filtered using the temporal signal space separation (tSSS) method [15-16] (Maxfilter, Elekta Neuromag Oy, Helsinki, Finland), which transforms the signals into magnetic multipole moments in terms of harmonic function amplitudes, separates into the internal and external subspaces divided by the sensor array, then suppresses the artifacts by omitting the harmonic function components of the external origin. The artifacts and interferences nearby the sensor array were further projected out when higher correlation coefficients ($p>0.98$) were obtained between the PCA (Principle Component Analysis) components in the temporal subspaces of 5 seconds block. The signals were then bandpass filtered using a fourth order bidirectional Butterworth filter with cutoff frequencies centered at f_m +/- 6 Hz/3 Hz width, averaged for 150 epochs and baseline-corrected using the signals of 1.25-seconds prestimulus period. Finally, the PCA components decomposed from the prestimulus period was projected out as noise.

Equivalent current dipole (ECD) was estimated for the stimulus presentation period by a single moving dipole method with a spherical volume conductor model using all 61 sensors that covered the left or right hemispheres, respectively (Curry 6.0.20, Compumedics Neuroscan, El Paso, TX, USA). The estimated ECD parameters at each time sample were evaluated in terms of Goodness of Fit (> 80 %) and 95 % Confidence Volume (< 1000 mm^3). The ECD result that satisfied all the terms and showed the most optimal parameters was identified to serve as a template model representing a source origin of the ASSR per hemisphere per subject.

The time course of the ECD moment, i.e., source waveform, was calculated by applying the template model to the Multiple Signal Classification (MUSIC) method [17], in which the signal subspace orthogonal to the noise space was associated with the ICA (independent component analysis)-filtered leading singular values of the measured data matrix. The maximum ECD moment was defined for each chord. Further, the time axis of the source waveform obtained for the SAM chirp tones was transformed into the frequency axis according to the sweeping f_c of the stimulus tones. The instantaneous ECD moment was identified on the frequency axis at seven discrete f_c corresponding to f_2 of each chord, to represent the intrinsic frequency characteristics of the ASSR. The results were normalized and evaluated statistically using a three-ways repeated measures ANOVA with factors of hemisphere (2), chords vs. SAM chirp tone (2) and f_c (7). Further, the effect of frequency characteristics of the ASSR was compensated by utilizing the results from the SAM chirp tone and tested using a two-ways repeated measures ANOVA with factors of hemisphere (2) and frequency ratio (7) and a post-hoc Scheffe's test.

III. Results

Figure 2 exemplifies the MEG signals at each stage of the analysis procedure. The temporal courses of the raw, tSSS filtered and bandpass filtered signals of one sensor from right temporal area of one subject are shown in (a), (b) and (c), respectively. The fast-Fourier transformed spectra are shown on their right column, displaying the ASSR component clearly elicited at 41.6 Hz in the spectral domain. The 150-epochs averaged signal is shown in (d), confirming also in the temporal domain the augmentation of the ASSR component during the stimulus presentation period. The source location of the ASSR was estimated in around the auditory cortices as displayed in (e). The time course of the source activity originated from the source location is shown in (f) for a chord, together with the signal from the lower frequency range (1-40 Hz) overlaid to show the evoked responses, P1m, N1m and P2m components, for reference (left), and for an ascending SAM chirp tone (right). It was observed that the ASSR to the chord started to increase after the stimulus onset and reached to the steady state at around 300 ms when the evoked responses were settled to the baseline.

Figure 3 (a) plots the strength of the ECD moment for the right hemisphere in response to the chords (black like) and to the SAM chirp tone (gray line). The f_2 of the chords corresponds to the instantaneous f_c of the SAM chirp tone. Both curves showed a monotonic increase along with f_c to peak at around 500 Hz and a decrease towards 624 Hz. The effect of f_c was significant ($F(6,24)=3.00$; $p<0.03$), confirming that the strength of the ECD moment of the ASSR was varied as a function of carrier frequency. There was also an interaction effect of f_c by chord vs. SAM chirp tone ($F(6,24)=6.18$; $p<0.01$), with the chords eliciting significantly smaller response than the SAM chirp tones at 128 Hz ($p<0.01$).

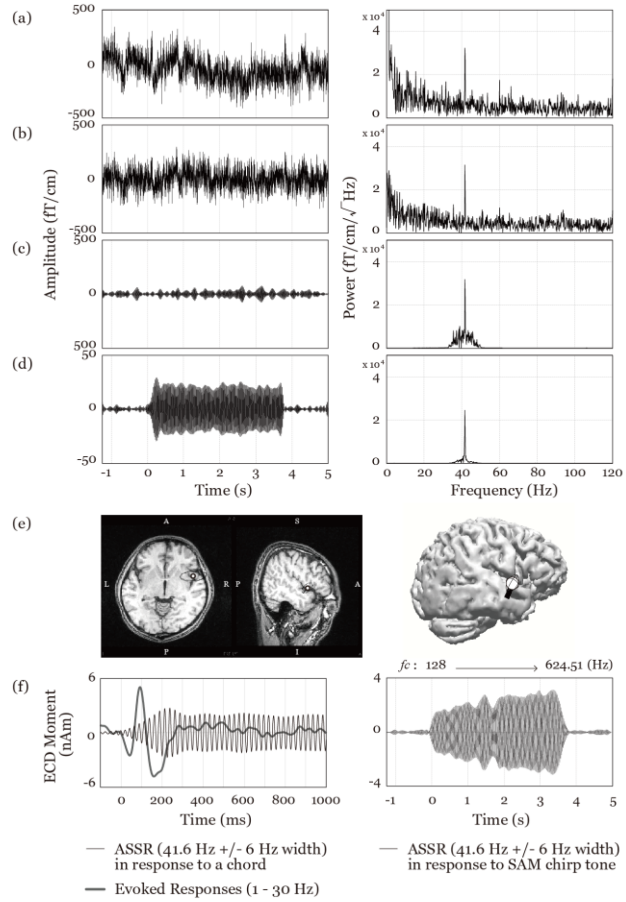

Figure 2. MEG signal analysis procedure. (a)-(d) Raw, tSSS-filtered, bandpass filtered and 150-trials averaged signals from one sensor in temporal (left column) and spectral (right column) domains, respectively. (e) Result of ECD estimation overlaid on MR and 3D reconstructed images. (f) The time course of the ASSR originated from the source location for a chord, together with the signal from the lower frequency range (1-30 Hz) showing the evoked responses, P1m, N1m and P2m components, overlaid for reference (left), and for an ascending SAM chirp tone (right).

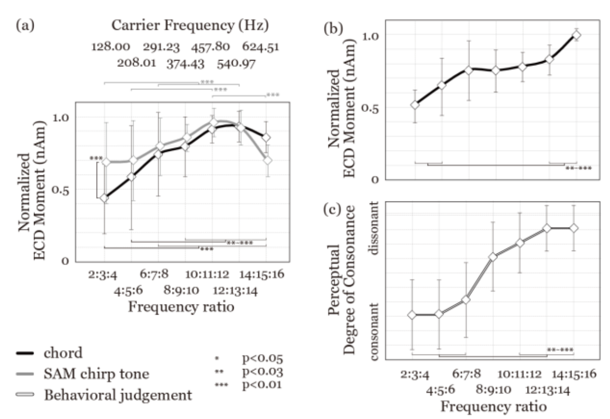

Figure 3. (a) The strength of the ECD moment obtained from the right hemisphere in response to chords (black line) and SAM chirp tone (gray line). (b) Frequency characteristics intrinsic to ASSR were compensated from the chords. (c) Behavioral judgment for perceptual degree of consonance.

Figure 3 (b) shows the strength of the ECD moment when the effect of the frequency characteristics of the ASSR was compensated by utilizing the responses to the SAM chirp tones. Significant difference was observed between the chords ($F(6,24)=8.85$; $p<0.01$), with the chords of simple frequency ratios (2:3:4, 4:5:6) eliciting smaller responses than the chords of complex frequency ratios (12:13:14, 14:15:16) ($p<0.01$~0.03). In the behavioral judgment, the chords of simple frequency ratios were evaluated with the higher degree of consonance whereas the chords of complex frequency ratios were perceived as dissonant as shown in (c). The difference was statistically significant ($F(6,24)=2.70$; $p<0.04$), with 2:3:4, 4:5:6 and 6:7:8 being more consonant than 10:11:12, 12:13:14 and 14:15:16 ($p<0.01$~0.03).

IV. Discussion

In the present study, we compared the strength of the ASSR phase-locked to Δf of the chords where the frequency ratio between the component tones was systematically varied. In agreement with the earlier study [8], the ASSR showed the significant difference between the chords, with those of simple frequency ratio being smaller than those of complex frequency ratio. In the present study, the effects of the acoustic parameters of the chords and the intrinsic neuronal characteristics of the ASSR were under control. That is, the frequency as well as the spectral energy of Δf was physically equal at 41.6 Hz in all the chords. The chords were also delivered at the equal sound intensity level with a deviation of +/- 0.5 dB SPL. Further, the responsiveness of the auditory cortical neurons which varies as a function of carrier frequency was compensated by utilizing the frequency characteristics obtained in response to the SAM chirp tone [13]. Nevertheless, the ASSR showed the significant difference between the chords, suggesting that the modulation of the ASSR was not attributed to the acoustic parameters of the Δf.

At the auditory peripheral loci, the neuronal activities represent more faithfully the acoustic features of the stimulus sounds. However, through the successive processing along with the afferent pathway, information is integrated by converging input and transformed into internal representation at the cortical level [18]. The encoding process is largely associated with neuronal preference for particular feature of input. Wang et al (2005) showed that the responsiveness of the auditory neurons at the onset of the input progressively decreases for preferred stimuli whereas it stays constant through to the cortex for non-preferred stimuli [18-19]. It has been suggested that the ASSR is a superposition of the mid-latency onset-evoked response (MLR) originating from the auditory cortices [5,10,11]. It is therefore considered that the results of the present study might represent the neuronal behavior at the onset detection where the simple frequency ratio was processed as preferred and the complex frequency ratio as non-preferred input. It also agrees with the attentional gating characteristics of the thalamo-cortical function [20]. It is to note that the other neurophysiological components such as sustained response and spontaneous rhythmic activities might show different behavior.

V. Conclusion

The present study explored how the strength of the neuromagnetic ASSR would be varied as a function of the frequency ratio of chords. The ASSR showed the significant difference between the chords, with those of simple frequency ratio elicited smaller response than those of complex frequency ratio did. It might represent the neuronal systematic behavior at the stage of onset detection where the simple frequency ratio was recognized as preferred/natural/standard information thus processed in energy saving manner whereas the complex frequency ratio was processed as non-preferred/artificial/deviant input without suppression.

Acknowledgment

The study was supported by the Funding Program for Next-Generation World-Leading Researchers provided by the Cabinet Office, Government of Japan.

References

[1] H. L. F. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, Trans. A. J. Ellis. New York: Dover, 1964.

[2] R. Plomp and W. J. Levelt, "Tonal consonance and critical bandwidth", *J Acoust Soc Am*. 38, 1965, 548-60.

[3] R. Plomp and H. J. Steeneken, "Interference between two simple tones", *J Acoust Soc Am*. 43, 1968, 883-4.

[4] P. X. Joris *et al*., "Neural processing of amplitude-modulated sounds", *Physiol Rev*. 84, 2004, 541-77.

[5] T. W. Picton *et al*., "Human auditory steady-state responses", *Int J Audiol.* 42, 2003, 177-219.

[6] D. R. Stapells et al., "Human auditory steady state potentials," *Ear Hear*. 5, 1984, 105-13.

[7] G. Langner, "Periodicity coding in the auditory system," *Hear Res.* 60, 1992, 115-42.

[8] Y. I. Fishman *et al*., "Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans," *J Neurophysiol*. 86, 2001, 2761-88.

[9] M. Ménard *et al*., "Relationship between loudness growth function and auditory steady-state response in normal-hearing subjects", *Hear Res.* 235, 2008, 105-13.

[10] R. Galambos et al., "A 40-Hz auditory potential recorded from the human scalp," *Proc Natl Acad Sci USA*. 78, 1981, 2643-7.

[11] B. Ross B et al., "A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones," *J Acoust Soc Am*. 108, 2000, 679-91.

[12] C. Wienbruch et al., "Frequency organization of the 40-Hz auditory steady-state response in normal hearing and in tinnitus," *Neuroimage*. 33, 2006, 180-94.

[13] A. Otsuka *et al*., "Frequency Characteristics of Neuromagnetic Auditory Steady-State Response (ASSR)," *Proc Bio Phys Eng*, 2012, 219-224.

[14] R Hari et al., "Neuromagnetic steady-state responses to auditory stimuli," *J Acoust Soc Am*. 86, 1989, 1033-9.

[15] S. Taulu et al., "Applications of the signal space separation method," *IEEE Trans Sign Proc*. 53, 2005, 3359–3372.

[16] S. Taulu and J. Simola, "Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements," *Phys Med Biol*. 51, 2006, 1759-68.

[17] JC. Mosher et al., "Multiple dipole modeling and localization from spatio-temporal MEG data," *IEEE Trans Biomed Eng*. 39, 1992, 541-57.

[18] X. Wang, "Neural coding strategies in auditory cortex," *Hear Res*. 229, 2007, 81-93.

[19] X. Wang et al., "Sustained firing in auditory cortex evoked by preferred stimuli," *Nature*. 435, 2005, 341-6.

[20] M. H. Giard et al., "Neurophysiological mechanisms of auditory selective attention in humans," *Front Biosci*. 5, 2000, D84-94.