31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Auditory MEG Responses to Removal of a Tone in C Major Scale Measured in a Pair of Reciprocal Oddball Schemes

Iku Nemoto, *Member, IEEE* and Masahiro Shoji

***Abstract*—The C major scale was used either as frequent or as infrequent stimulus in the oddball auditory evoked field measurement where the other stimulus was constructed by removing one the tones in the scale. Multivariate statistical analysis was employed to judge whether there was a significant difference between the responses to complete and incomplete scales in *each* subject for each 'target' tone which was removed in an incomplete scale. Incomplete scales lacking, especially E, or B caused responses in both of the two oddball schemes but less significantly when used as frequent stimuli indicating that the complete major scale stored in the long term memory retained its influence as 'reference' stimulus even when presented with a smaller probability.**

I. Introduction

Musical scale is one of the fundamentals in western music. As western music has spread worldwide, so have the major and minor scales which have been the most important scales in western music for centuries. It would be difficult to determine whether some inherent physiological structures in the brain are responsible for the worldwide acceptance of western music and of these scales in particular. Instead we can search in the brain for imprints of musical experiences of individuals by means of EEG, MEG and fMRI. Many of the studies of musical perception by these techniques [1-12] have focused on responses to incongruities in fragments of music constructed for particular experiments. The incongruities have been those in chord sequences or in tonal contexts, for example, and the results have indicated that individual musical experiences have organized structures in the brain which reflect those important aspects in western music which have been described in music theory.

However, we have felt that the fragments of music used as stimuli in the previous studies were somewhat arbitrary and the whole experimental design was not sufficiently systematic. Naturally, it is impossible to prepare 'all' combinations of tones to perform an exhaustive study to be systematic. Instead, one could restrict the stimuli within a well-defined class and do an exhaustive study. The present study could be a starting point of that kind of study. We used the complete C major scale (one octave, from C to C) and incomplete ones each of which lacked one tone of the scale. Previously, we used two kinds of incomplete scale; in PBS (prolong before skip), the tone preceding the one skipped was prolonged and in RBS (repeat before skip) the preceding tone was repeated (Fig. 1). We used the complete scale as frequent (or standard) stimulus and one of the incomplete scales as infrequent (or deviant) stimulus and measured the difference in the auditory evoked magnetic field (AEF). This procedure is often called an oddball paradigm. The result showed that in both PBS and RBS experiments, removal of the tones E, G and B elicited large difference (mismatch field, MMF). One question remained as regards the mismatch field. Usually, mismatch field (or potential in case of EEG) is defined to be the field overlapping the N1m response (peaking around 100 ms after the onset of the tone), peaks at around 130-200 ms after the onset of the infrequent stimulus and is supposed to be caused by the novelty of the infrequent stimulus. The novelty is supposed to come from comparison of the incoming stimulus with the frequent stimulus which is temporarily stored in the short term memory. However, the complete scale was familiar to all the subjects who were no musicians but had musical education on the comparison-education level in elementary and junior high schools (age 6-19) and heard western-style music practically everyday. Therefore, although the C major scale was temporarily stored in the short term memory, it must have been stored in the long term memory also. So what we treated as mismatch field may have been different from the usual definition of the term and closer to what has been termed ERAN (early right-anterior negativity) [8].

In the present study, we presented RBS scales more frequently than the complete scale (call this ROB, reverse oddball) and compared the result of the experiment (call this NOB, normal oddball) with frequently occurring complete scale. We avoid the terms 'standard' and 'deviant' as they

(a) C D E F G A B C

(b) C D D F G A B C

Fig. 1 The complete C major scale (a) and RBS-E (b), an incomplete scale in which the 'target' tone E has been replaced by repeated D.

Manuscript received April 6, 2009. This work was supported in part by the the Ministry of Education, Science, Sports and Culture, Japan.

Iku Nemoto is with the Department of Information Environment, Tokyo Denki University, Muzai-gakuendai, Inzai, Chiba, 270-1382, Japan (phone & fax: +81-476-46-8424; email: nemoto@sie.dendai.ac.jp).

Masatoshi Shoji was with the same department when this work was done.

may cause confusion. The purpose of reversing the probabilities between the two scales was to see whether the infrequent stimulus (complete scale in ROB) would elicit a mismatch field as the usual oddball paradigm would. Or would incomplete scales still elicit responses to 'novelty' because the complete scale is strongly fixed as reference? The result was more in favor of the second possibility but the responses were smaller suggesting that frequent incomplete scales temporarily weakened the 'normality' of the complete scale.

Because responses to musical stimuli in general greatly vary among subjects, individual responses should be focused on as well as group data. For this purpose, we applied multivariate analysis to sample vectors made of raw (unaveraged) data of each subject.

II. Methods

A. Subjects

Thirteen male students (age 21 or 22) of Tokyo Denki University with no record of hearing abnormalities participated in the NOB experiments and 9 of them also participated in the ROB experiments. Experiments were done in accordance with the Helsinki Declaration and were approved by the ethical committee of the university. Written forms of informed consent were obtained from the subjects. None of them had professional musical training but had musical education on the compulsory-education level (age 6-15). One of the 9 subjects taking part in ROB experiments took private piano lessons for 5 years as a child.

B. Stimuli

The complete C major scale started with C at $440 \times 2^{-3/4}$ Hz and ended with C at $440 \times 2^{1/4}$ Hz. Each tone written as an eighth note in Fig. 1 lasted for 200 ms including 3 ms rise and fall times. An RBS scale was constructed by replacing one tone (called 'target' tone hereafter) by the tone immediately preceding it. See the example RBS-E in Fig. 1 (b). The intensity of every stimulus was 60 dB above the threshold of hearing for each subject measured immediately before each experimental session. The loudness change with the frequency was compensated for by using the Fletcher-Munson equal loudness curve. The sound was generated by ER-2 earspeaker, Etimo, USA and delivered to the subject's left ear via a plastic tube 1.7 m long.

C. Reciprocal Oddball Pair and MEG Measurement

The frequent and infrequent stimuli were presented in random order. Table 1 summarizes the paradigm. In NOB (normal oddball) procedure, the complete C major scale was presented with probability 0.8 and an incomplete scale with probability 0.2. In ROB procedures the probabilities were reversed. The inter stimulus interval (between scales) was fixed at 0.5 s. MEG was measured in a magnetically shielded room by Neuromag122. EOG (electro-oculogram) was measured simultaneously to record eye movements and blinkings. The stimuli continued until more than 80 artifact-free epochs (EOG amplitude < 150 μV, MEG amplitude < 3000 fT/cm) of infrequent stimulus were recorded. The sampling frequency was 512 Hz and the raw data were filtered by FIR bandpass (1-40 Hz) filter of 200-th order before statistical processing and baseline treatment was done to each epoch using the 100 ms interval preceding it.

D. Statistical Analysis of Responses of Each Subject

We directly applied multivariate variance analysis to individual data made from raw (unaveraged) data. We call the following 108-th order vector a response vector:

$$\begin{aligned} &\mathbf{x} = (x_1, \ldots, x_k, \ldots, x_{108}) \\ &x_k = x_i(t_j),\ k = (i-1)*J + j,\ i = 1,\ldots,I \qquad (1) \\ &t_j = 80 + 25(j-1),\ j = 1,\ldots,J,\quad I = 12,\ J = 9 \end{aligned}$$

where the time t_j is measured with respect to the first point of the 200 ms time interval corresponding to the target tone, $x_i(t_j)$ is the raw measurement data of the $i-$th sensor at time t_j. The 12 sensor coils chosen for analysis were those roughly covering the auditory field of the right hemisphere. The time increment between successive elements was set at 25 ms as in the above. Let $\mathbf{x}_n, n = 1,2,\ldots,N_S$ be the vectors in response to the frequent stimulus formed as eq.(1) and $N_S > 300$ be the number of frequent-stimulus epochs. Also $\mathbf{y}_n, n = 1,2,\ldots,N_D$ are response vectors to the infrequent stimulus and $N_D > 80$ is the number of *infrequent* epochs. The sample $\mathbf{x}_n, n = 1,2,\ldots,N_S$ was assumed to be a random sample of size N_S from a q-variate (here we use $q = 108$) normal population with mean $\mathbf{m}_S$ and covariance matrix Σ_S, those for the infrequent stimulus being denoted by $\mathbf{m}_D$ and Σ_D. As we cannot hope for a large sample size N_D compared to the vector size $q = 108$, we also assume that $\Sigma_D = \Sigma_S$. To test the null hypothesis $\mathbf{m}_S = \mathbf{m}_D$, we calculate the test statistic [13]:

$$F = \frac{(N_S + N_D - q - 1)N_S N_D}{q(N_S + N_D)} (\bar{\mathbf{y}} - \bar{\mathbf{x}}) V^{-1} (\bar{\mathbf{y}} - \bar{\mathbf{x}})' \qquad (2)$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are sample mean vectors of $\mathbf{x}_n$ and $\mathbf{y}_n$ respectively, and

$$\begin{aligned} &V = V_S + V_D \\ &V_S = \sum_1^{N_S} (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^t,\ V_D = \sum_1^{N_D} (\mathbf{y}_n - \bar{\mathbf{y}})(\mathbf{y}_n - \bar{\mathbf{y}})^t \end{aligned}$$

Under the null hypothesis, the statistic F is distributed according to the F distribution with degrees of freedom, $\phi_1 = q, \phi_2 = N_S + N_D - q - 1$. If the F test rejected the null hypothesis, we would conclude that there was a siginificant mismatch field.

The same response vectors were used to judge whether there was a significant mismatch field in the group data by employing 2-way layout MANOVA, subjects and scales (complete and incomplete) being used as two factors.

III. Results

Fig. 2 shows the responses shown by one subject S18 to frequent and infrequent stimuli in NOB (a) and ROB (b) procedures. RMS values of the average measurements from 12 sensor coils are shown. In both of the procedures, the 'target' tone was A. It is seen that regardless of whether it was the frequent or infrequent stimulus, the incomplete scale showed larger responses although the amplitude became smaller. It is also noted that there were two peaks, the first one in response to the repeated G (instead of A) and the second one to B. The second one probably responded to the jump from G to B but we paied attention primarily to the first peak.

Fig. 3 shows the 'mismatch' responses of all the subjects. Each trace is for each target tone and is the rms across the subjects. In (a) the results of 13 subjects in NOB runs are shown and (b) shows the results of 9 subjects in ROB runs. The latter shows smaller responses than the former. The two-way MANOVA applied to the group data for NOB results (a) revealed that all the mismatch responses were significant ($p < 0.01$) and there were significant interactions ($p < 0.01$) between subjects and the stimuli. ROB results (b) showed significant responses to the removal of E, G, A and B. D and F did not elicit significant mismatch fields.

Table 1 shows the significance probability p for each tone (column) and each of those subjects (row) who participated in both NOB and ROB experiments. N in the parentheses after the tone name stands for NOB and R for ROB experiments. The asterisks are for cells with $p < 10^{-4}$.

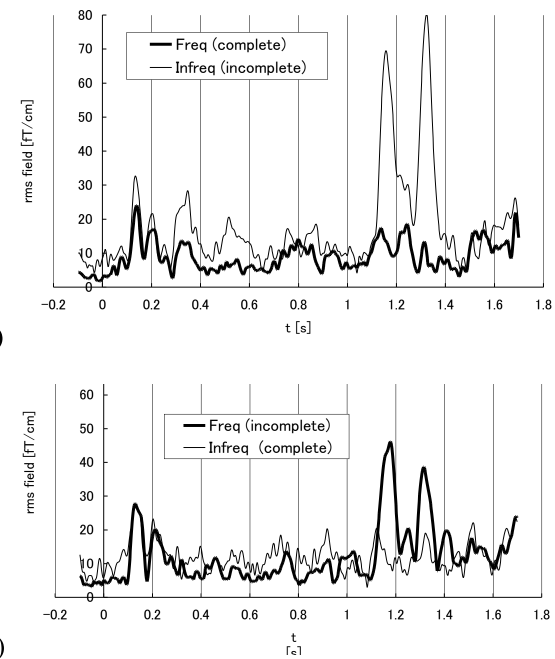
IV. Discussion

A. Mismatch fields?

If one compares the response to an infrequent stimulus with that to a frequent stimulus, the former has to be larger when there is a '*mismatch* field' as usually defined and mentioned in the Introduction. Fig. 2 (a) seems to fit this definition but the waveforms in (b) may refute this interpretation. Here, the incomplete scale presented more frequently than the complete scale still showed a large response. The 'novelty' therefore was not in the infrequent complete scale but was still in the incomplete scale despite its more frequent presentation. Therefore, the mismatch-like field shown in (a) may not be a mismatch field in the above narrow sense, either. The very big response to the incomplete scale in (a) may be caused by the novelty with respect to the complete scale stored in the long term memory and reinforced by frequent presentation.

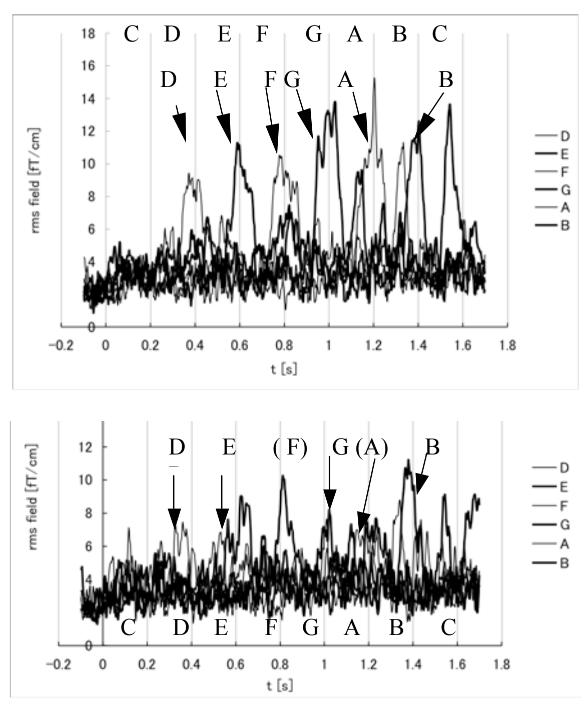
Table 1. The experimental scheme.

	Frequent Stimulus	Infrequent Stimulus
NOB	Complete Scale	Incomplete Scale Lacking 1 Tone
ROB	Incomplete Scale Lacking 1 Tone	Complete Scale



(a) (b)

Fig. 2. Results of one subject to complete C major scale and to the incomplete scale lacking A. (a)NOB; complete scale was frequent. (b)ROB; incomplete scale was frequent.



(a) (b)

Fig3. The results of NOB (a) and ROB (b) experiments for all the subjects and for all the target tones. See the text.

Table 2. The significance probability p for each target tone and subject in NOB (N) and ROB (R) runs. The asterisks are for the p values less than 10^{-4}.

	D (N)	E (N)	F (N)	G (N)	A (N)	B (N)	D (R)	.E (R)	F (R)	G (R)	A (R)	B (R)
S4	0.2131	0.0424	0.03771	0.5279	0.1836	**0.0002**	0.6911	**0.0015**	1.0000	0.1663	0.2912	0.5825
S6	0.1041	0.2222	**0.0006**	**0.0009**	**0.0014**	**0.0041**	0.0848	**0.0063**	0.1565	**0.0027**	0.3880	0.1720
S12	*	**0.0011**	0.1856	**0.0056**	**0.0005**	*	0.1247	0.2948	0.2928	0.1939	0.0904	0.3014
S13	0.9721	**0.0036**	0.1871	**0.0004**	0.023	*	0.9932	**0.0058**	0.3008	0.2003	0.4155	0.1379
S14	0.3208	0.6848	0.5836	**0.0035**	0.1063	0.7702	0.2073	0.3371	0.8330	0.3971	0.0639	0.0898
S15	0.0523	*	*	*	*	*	0.8521	0.2755	0.4427	0.0125	0.0512	**0.0004**
S16	0.0120	0.0230	0.3548	0.3817	*	**0.0003**	0.8168	0.1151	0.0964	0.0766	0.2714	**0.0052**
S18	*	*	**0.0001**	*	*	*	0.0722	**0.0008**	0.3501	0.5115	**0.0001**	**0.0015**
S19	*	*	*	*	*	0.1995	0.1356	**0.0012**	0.3501	0.0656	*	**0.0016**

In (b) the reversed probabilities 'weakened' the novelty of the incomplete scale but did not 'reverse' the 'normality' relationship between the two scales.

B. Tonic Triad and Leading Note

Fig. 3 (a) shows larger responses for the removal of E, G, B or A than of D or F. In the ROB experiment (b), the removal of E or B made large responses although the amplitudes were smaller than in (a). The tones E and G, together with C, make up the most important chord in C major, tonic triad, and the tone B is the leading tone which leads a melody to the tonic tone C. Therefore, the large responses to E, G and B seem to agree with the importance of these tones in the music theory. The importance of the leading tone has not been highlighted in the EEG or MEG literature before. The large responses to the removal of the tone A in NOB using the RBS (repeat before skip) type scale are yet to be explained.

C. Comparison among Subjects

The application of multivariate analysis to individual raw data made it possible to determine whether mismatch fields to removal of target tones were significant or not for each subject as seen in Table 1. For example, subject S19 showed significant mismatch fields to removal of 5 tones in NOB and to that of 3 tones in ROB. Subject S12 responded significantly to 5 tones in NOB also but no significant responses were observed in ROB. This indicates that S19 showed greater influence of the complete scale stored in the long term memory than S12 did. It would be interesting to see responses of those who have little been exposed to western music, but subjects would be hard to find. We now look at the number of subjects who showed significant 'mismatch' responses to the removal of each target tone. Let X(a, b) mean that a subjects showed significant mismatch response to the target X in NOB experiment and b subjects in ROB. We read from Table 1, D(3, 0), E(5, 5), F(4, 0), G (7, 1), A(6, 2) and B(7, 4). Here B and E seem to stand out. It may be an effectvie method in general to discuss the number of subjects who show siginificant responses such as done here in addition to discuss the significance of responses of the group data.

V. Conclusion

Incomplete scales laking one note of the C major scale caused a mismatch-field like response even when presented more frequently than the complete scale although it was smaller than when presented less frequently, indicating a strong influence of the major scale stored in the long-term memory. The results depended largely on the removed tones which was in agreement with the musical theory. Individual differences were also well demonstrated.

References

[1] R. J. Zatorre, A. C. Evans and E. Meyer, "Neural mechanisms underlying melodic perception and memory for pitch," *J. Neurosci,* vol. 14, pp.1906-1919, April 1994.

[2] E. Granot, "Do Re Mi Fa Sol La Ti – Constraints, Congruity, and Musical Training: an event-related brain potentials study of musical expectancies," *Music Perception*, vol. 19, pp. 487-528, June 2002.

[3] L. Minati, C. Rosazza, L. D'Incerti, E. Pietrocini, L. Valentini L, Scaioli V, et al. "FMRI/ERP of musical syntax: comparison of melodies and unstructured note sequences," *NeuroReport*, vol. 19, pp.1381-1385, Sept. 2008.

[4] A. H. Foss, E. L. Altschulter and K. H. James, "Neural correlates of the Pythagorean ratio rules," *NeuroReport*, 2007; vol. 18, pp. 1521-1525, Oct. 2007.

[5] T. Mizuno and M. Sugishita, "Neural correlates underlying perception of tonality-related emotional contents," *NeuroReport*, vol.18, Oct. 1651-1655, 2007.

[6] S. Kuriki, N. Isahai and A. Ohtsuka, "Spatiotemporal characteristics of the neural activities processing consonant/dissonant tones in melody," *Exp Brain Res*, vol. 162, pp. 46-55, Mar. 2005.

[7] B. Maess B and S. Koelsch, T. C. Gunter and A. D. Friederici, Musical syntax is processed in Broca's area: an MEG study, *Nature Neurosci*, vol. 4, pp. 540-545, May 2001.

[8] S. Koelsch, T. C. Gunter, E. Schroeger, M. Tervaniemi, D. Sammler, A. and D. Friederici, "Differentiating ERAN and MMN: an ERP study," *NeuroReport*, vol. 12, 1385-1389, May 2001.

[9] S. Koelsch, T. C. Gunter and A. D. Friederici, "Brain indices of music processing: Nonmusicians are musical," *J Cogn. Neurosci.,* vol. 12, pp. 520-541, May 2000.

[10] S. Koelsch, E. Schroeger and M. Tervaniemi, Superior pre-attentive auditory processing in musicians, *NeuroReport*, vol. 10, 1309-1313, April 1999.

[11] T. Hashimoto, Y. Hirata and S. Kuriki, "Auditory cortex responds in 100 ms to incongruity of melody," *NeuroReport*, vol. 11, pp. 2799-2801, Aug. 2000.

[12] P. Janata, J. J. Birk, J. D. Van Horn, M. Leeman, B. Tillmann and J. J. Bharucha, "The cortical topography of tonal structures underlying Western music," *Science*, vol. 298, pp. 2167-2170, Sept 2002.

[13] R. A. Johnson and D. W. Wichern, "Applied Multivariabe Statistical Analysis," Prentice Hall, 2002.