Auditory cortex responds in 100 ms to incongruity of melody

Tatsuaki Hashimoto, Yoshihiro Hirata and Shinya Kuriki[CA]

Research Institute for Electronic Science, Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo 060-0812, Japan

[CA]Corresponding Author

Received 17 May 2000; accepted 18 June 2000

In order to find the cortical neural activities that reflect musical processing, we conducted measurements of MEG responses while subjects listened to familiar melodies. Significant enlargement of the N1m peak was observed in the response to the endnote of a phrase that was deviated to out-of-key from the keynote. Equivalent current sources of the enlarged N1m response were localized in the cortical area of the superior temporal plane. These results suggest the possibility of fast neural processing, as early as 100 ms, of the auditory cortex detecting incongruity of a melody. *NeuroReport* 11:2799–2801 © 2000 Lippincott Williams & Wilkins.

Key words: Auditory cortex; Magnetic fields; Magnetoencephalogram (MEG); Melody; Music; N1m

INTRODUCTION

A melody is a temporal sequence of pitches or harmonies which fit the scale of a selected key. Incongruity occurs when the pitch or harmony of notes deviates within a phrase, and this deviation elicits a late positive wave or P300 of the ERP (event-related potential) [1–3]. Although P300 can be an index of neural activity detecting musical incongruity, it is thought to reflect unexpectedness of a rare event without specificity. A sentence in a language is composed of sequential words that are syntactically organized. Incongruity of a sentence can be generated either semantically or syntactically by altering the words. Semantic and syntactic incongruities elicit distinct ERP components of N400 and left anterior negativity, respectively, with different scalp topographies [4,5]. We tried to determine the responses that would reflect musical processing by measuring MEG signals, which have high spatial selectivity around the scalp. Significant enlargement was observed in the N1m responses that were elicited by incongruent endnotes in phrases of familiar melodies. Equivalent sources of the N1m responses were localized in the auditory cortex.

MATERIALS AND METHODS

Seven normal hearing right-handed subjects with a mean age of 27.0 years (range, 21–34 years) participated in the experiment. Informed consent was obtained from each of the subjects after they had been given an explanation of the experiment. End phrases were taken from 25 well-known songs which the subjects had learned before they started high school. Piano tones sampled and stored in a digital memory were used as stimuli. The mean (s.d.) length of the phrases and the number of notes in the phrases were 6.9 (1.4) s and 18.0 (5.1), respectively. The endnotes of the phrases were modified out-of-key in different manners, and 150 deviant phrases, each with an incongruent note (IN), were composed. These IN phrases were mixed in random order with the same number of phrases that each had a congruent endnote (CN) in the original melodies (Fig. 1a). The loudness of the tone of the endnote of IN and CN phrases was 75–80 dB in sound pressure.

MEG recordings were conducted using a 19-channel SQUID magnetometer [6] over the right side of subjects' head in a 15 cm radius area centered at a location 7–10.5 cm superior and 5 cm posterior–3 cm anterior to the pre-auricular point. The mixed CN and IN phrases were presented to the contralateral (left) ear of subjects at 2.0 s intervals. In order to maintain their attention, the subjects were instructed to count the number of IN melodies and report it aurally at the intermission after recordings of 25–30 melodies. They were also instructed to close their eyes during the recording.

The MEG signals were divided into CN and IN responses and averaged with reference to the onset of the endnote tone. The averaged responses were then filtered digitally at 1–40 Hz, with a time-average during a period at 1.7–2.0 s after the onset serving as a baseline. Single equivalent current dipole (ECD) sources were estimated by calculating their coordinates and moments in the head frame of subjects and registered in their MR images. ECD sources that had goodness-of-fit values of $>90\%$ (or $<10\%$ variance of residual fields) were selected as reliable sources. In order to examine reproducibility of the responses, subjects underwent the same MEG measurements 2 months later under the same conditions.

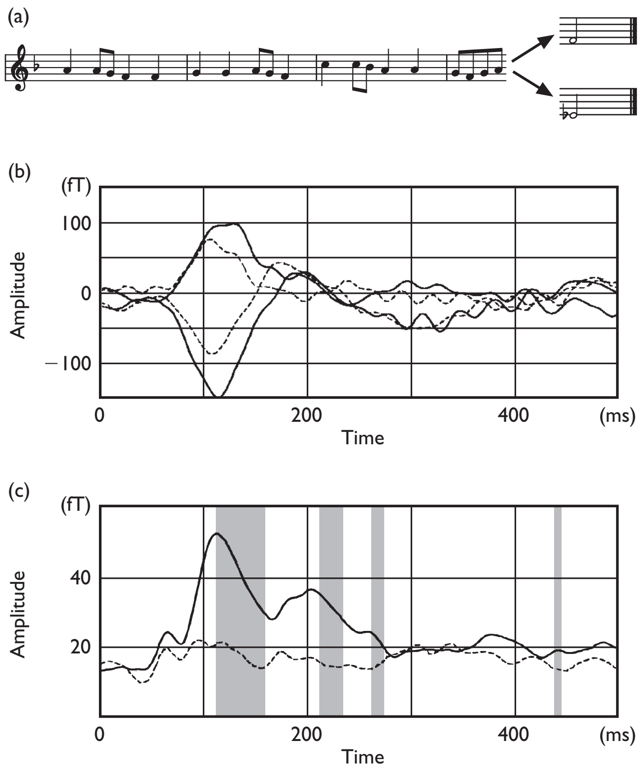

Fig. 1. (**a**) Example of a musical phrase ending with a congruent note (CN), i.e. keynote, (upper trace), and incongruent note (IN), i.e., out-of-key endnote, (lower trace). (**b**) Waveforms of the MEG responses recorded for the CN (broken line) and IN (solid line) in one subject. The waves of positive (outward) and negative (inward) maxima were selected from 19-channel responses. (**c**) Mean amplitude across subjects of root-mean-squared 19-channel responses elicited by the CN (broken line) and IN (solid line). Shadowed areas indicate significant difference ($p < 0.01$; paired t-test, $n = 7$) between the CN and IN responses.

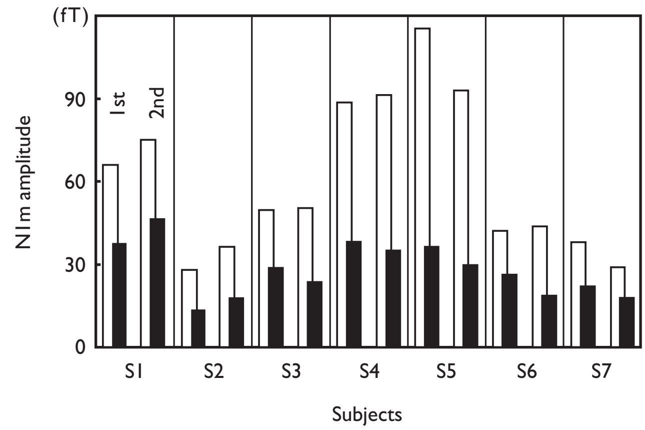

Fig. 2. Root-mean-squared amplitude of the N1m peak in the IN (white bar) and CN (shadowed bar) responses, which were obtained in the first and second measurements for individual subjects (S1–S7). A significant difference between the IN and CN responses for pooled data was confirmed ($p < 0.001$, paired t-test, $n = 14$).

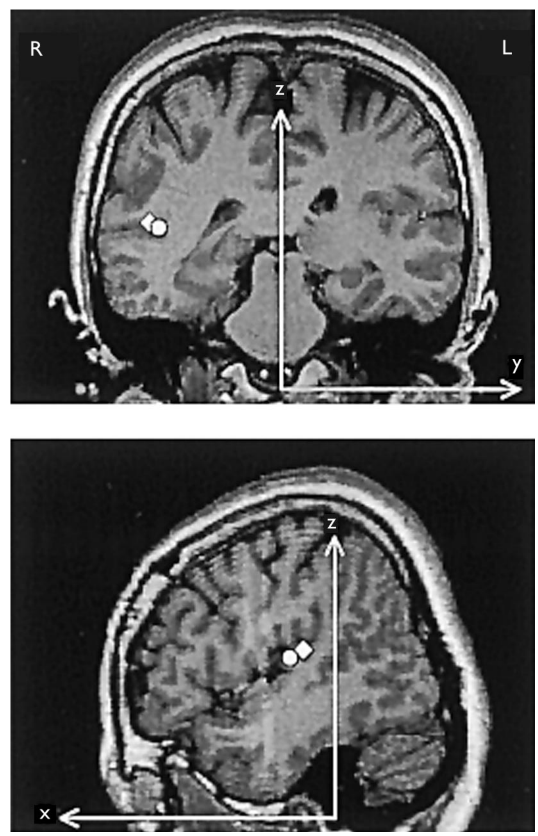

Fig. 3. Location of N1m sources for the IN (square) and first note (circle) responses obtained in one subject. The two sources were indistinguishably close to each other, both being located in the auditory cortex at the posterior part of the sylvian fissure.

RESULTS

Figure 1b shows waveforms of the MEG responses to the CN (broken line) and IN (solid line) tones for one subject. As in this case, a clear N1m peak at about 110 ms was observed for the IN response in all subjects. Positive (outward) and negative (inward) fields at the N1m peak indicate a dipolar distribution in the recording area, suggesting a single ECD source. The N1m peak in the CN response was decreased in amplitude or even unclear in some subjects. Mean amplitude across subjects of root-mean-squared 19-channel responses is shown in Fig. 1c. As indicated by the shadowed areas, the mean of the IN response (solid line) was significantly larger ($p < 0.01$; paired t-test, $n = 7$) than that of the CN response (broken line) during a period of 110–150 ms after the N1m peak and in later shorter periods. There was no significant difference in the latency of the N1m peaks in the IN (111 (28) ms) and CN (106 (23) ms) responses.

In the repeated measurements conducted 2 months later, very reproducible results of the N1m amplitudes were obtained for all subjects. Figure 2 shows the root-mean-squared amplitude of the N1m peak of the IN (white bar) and CN (shadowed bar) responses in the first and second measurements of individual subjects (S1–S7). A significant difference between the N1m amplitudes of the IN and CN responses for the pooled data was confirmed ($p < 0.001$, paired t-test, $n = 14$).

Reliable ECD sources at the latency of the N1m peak, with goodness-of-fit values of >0.9, were obtained for the IN responses in four of seven subjects. Such reliable ECD sources were not obtained for the CN responses in any of the subjects because of low amplitude. Additional localization of the N1m source was made for the responses, which were elicited by the tone of the first note in the phrase. In these first note responses, the N1m peak was commonly larger than that in the IN responses. Reliable ECD sources were obtained in all subjects. Figure 3 shows a comparison of the locations of the N1m sources for IN (square) and first note (circle) responses in one subject. The two sources were indistinguishably close to each other, both being located in the auditory cortex at the posterior part of the sylvian fissure. Similar results of closely located N1m sources for the IN and first note responses were obtained in all the four subjects for whom reliable sources were obtained.

DISCUSSION

Compared with the amplitude of the N1m response that was elicited by the first note tone of the phrase, the CN response was significantly attenuated in the process of continuing pitches. The enlargement of the IN response from the CN indicates partial recovery of this reduced activity, suggesting that the IN response reflects neural detection of the incongruity of the melody. The latency of N1m and the location of ECD sources further suggest a fast processing, as early as 100 ms, in the auditory cortex in the superior temporal plane. The N1m source generated by pure tones and speech sounds in the superior temporal plane is believed to arise from populational neural activities in the secondary or higher auditory area, posterior to the primary area of the Heshl's gyrus [7,8]. The location of the N1m source of the IN response, that was indistinguishable from the N1m source of the first note response, suggests that the cortical activity related to melody processing occurs in the secondary/higher auditory area, in agreement with the findings in a recent clinical study of music processing ability in patients [9].

The songs used in the experiment were simple and familiar to all subjects, and the IN in the melody was therefore highly unexpected. In previous EEG studies, late positive waves, such as P300, following endnotes that are incongruent in melody and harmony have been observed [1,3]. Here, the P300 wave is considered to reflect the unexpectedness of a rare event, i.e. an oddball, in the sequence of congruent stimuli [10]. However, the enlarged N1m wave in the IN response is distinct from such late potentials. A slightly enhanced negative potential at 100–200 ms has been reported for the incongruity of a melody, but this observation was not convinced by the authors [1]. The IN response is also distinct from the N400 wave of potential and its magnetic counterpart, which appear following semantic incongruities in auditory and visual sentences [4,5,11–13]. From the latency, the IN response is close to the anterior negativity that is elicited at about 180 ms by syntactically incorrect sentences [14], though the frontal source is estimated from the topography of the negativity.

CONCLUSION

We observed significant enlargement of the N1m peak in a response that was elicited by the endnote tone of a musical phrase deviated to out-of-key from the keynote. The equivalent current source of the enlarged N1m response was localized in a cortical area of the superior temporal plane. The source location was indistinguishable from that of the source of the response to the first note of the phrase. These results suggest that the posterior portion of the auditory cortex participates in the processing of a melody, in addition to the processing of acoustic characteristics of tones and speech sounds.

REFERENCES

1. Besson M and Macar F. *Psychophysiology* **24**, 14–25 (1987).
2. Paller KA, McCarthy G and Wood CC. *Psychophysiology* **29**, 202–206 (1992).
3. Besson M, Faïta F and Requin J. *J Neurosci Lett* **168**, 101–105 (1994).
4. Kutas M and Hillyard SA. *Science* **207**, 203–205 (1980).
5. Kutas M and Hillyard SA. *Nature* **307**, 161–163 (1984).
6. Hirata Y and Kuriki S. *IEICE Trans Electron* **E79-C**, 1213–1218 (1996).
7. Pantev C, Bertrand O, Eulitz C *et al. Electroencephalogr Clin Neurophysiol* **94**, 26–40 (1995).
8. Kuriki S, Okita Y and Hirata Y. *Exp Brain Res* **104**, 144–152 (1995).
9. Liégeois-Chauvel C, Peretz I, Babaï M *et al. Brain* **121**, 1853–1867 (1998).
10. Dunean-Johnson CC and Donchin E. *Psychophysiology* **14**, 456–467 (1977).
11. McCallum WC, Farmer SF and Pocock PV. *Electroencephalogr Clin Neurophysiol* **59**, 477–488 (1984).
12. Connolly JF and Phillips NA. *J Cogn Neurosci* **6**, 256–266 (1994).
13. Helenius P, Salmelin R and Service E. *Brain*, **121**, 1133–1142 (1998).
14. Friederici AD, Pfeifer E and Hahne A. *Cogn Brain Res* **1**, 183–192 (1993).