

Locating Melody Processing Activity in Auditory Cortex with Magnetoencephalography

Roy D. Patterson, Martin Andermann, Stefan Uppenkamp and André Rupp

Abstract This paper describes a technique for isolating the brain activity associated with melodic pitch processing. The magnetoencephalographic (MEG) response to a four note, diatonic melody built of French horn notes, is contrasted with the response to a control sequence containing four identical, “tonic” notes. The transient response (TR) to the first note of each bar is dominated by energy-onset activity; the melody processing is observed by contrasting the TRs to the remaining melodic and tonic notes of the bar (2–4). They have uniform shape within a tonic or melodic sequence which makes it possible to fit a 4-dipole model and show that there are two sources in each hemisphere—a melody source in the anterior part of Heschl’s gyrus (HG) and an onset source about 10 mm posterior to it, in planum temporale (PT). The N1m to the initial note has a short latency and the same magnitude for the tonic and the melodic sequences. The melody activity is distinguished by the relative sizes of the N1m and P2m components of the TRs to notes 2–4. In the anterior source a given note elicits a much larger N1m-P2m complex with a shorter latency when it is part of a melodic sequence. This study shows how to isolate the N1m, energy-onset response in PT, and produce a clean melody response in the anterior part of auditory cortex (HG).

R. D. Patterson (✉)

Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

e-mail: rdp1@cam.ac.uk

M. Andermann · A. Rupp

Section of Biomagnetism, Department of Neurology, University of Heidelberg, Heidelberg, Germany

e-mail: andermann@uni-heidelberg.de

A. Rupp

e-mail: andre.rupp@uni-heidelberg.de

S. Uppenkamp

Medizinische Physik and Cluster of Excellence Hearing4All, Carl von Ossietzky Universität Oldenburg, 26111 Oldenburg, Germany

e-mail: stefan.uppenkamp@uni-oldenburg.de

© The Author(s) 2016

P. van Dijk et al. (eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, Advances in Experimental Medicine and Biology 894, DOI 10.1007/978-3-319-25474-6_38

Keywords Auditory cortex · Melodic pitch · Magnetoencephalography

1 Introduction

Recently, Patterson et al. (2015) extended earlier imaging research on melody processing (Patterson et al. 2002; Rupp and Uppenkamp 2005; Uppenkamp and Rupp 2005) by presenting listeners with four-bar musical phrases played by a French horn. There were four notes per bar; they had the same amplitude and the inter-note gaps were minimized to minimize the energy onset component of the responses to the second and succeeding notes 2–4 (Krumbholz et al. 2003). In the first and third bars, the pitch was restricted to the tonic of the four-bar phrase; the second bar contained a novel, random tonal melody and the fourth bar contained a variation of that melody. Magnetoencephalographic recordings revealed that each note produced a prominent transient response (TR) in auditory cortex with a P1m, an N1m and a P2m. The structure of the four-bar phrase was reflected in the magnitude of the TR. In all bars, the TR to the first note was somewhat larger than that to the remaining three notes of that bar. In the melody bars, the P2m component of the TR was larger than in bars where the pitch was fixed. The stability of the TR magnitude to notes 2–4 of each bar made it possible to do a four-dipole analysis and show that the activity is distributed between two sources in each hemisphere, one in the medial part of Heschl's gyrus (HG) and the other slightly more lateral and posterior on the front edge of planum temporale (PT). The source waves showed that melody processing was largely confined to the P2m component of the TR in the medial part of HG.

This paper describes a shorter version of the melody experiment with alternating tonic and melodic bars and an inter-sequence interval of about 1325 ms—modifications that make it possible to measure the sustained field (SF) generated during the tonic and melodic bars. As in the previous experiment, the magnitude of the TR to the first note of any sequence is enhanced by the abrupt onset of energy. The purpose of this second experiment was to determine (1) whether the TR to the first note of each sequence emanates from a different source than the TRs from the remaining notes of that sequence, and (2) whether the source of the SF is in auditory cortex or beyond.

2 Methods

A natural French horn tone (Goto et al. 2003) served as the basic stimulus for the experiment. Four-note sequences were constructed by varying the f_0 of the original note within the original envelope using the vocoder STRAIGHT (Kawahara et al. 1999; Kawahara and Irino 2004). There were two different types of four-note sequences, “tonic” and “melodic”: The f_0 was fixed in the tonic sequences at either 100 Hz or 141.5 Hz. In the melodic sequences, the f_0 value of each note was

randomly drawn, without replacement, to be -1 , 0 , 2 , 4 , 5 or 7 semitones relative to the respective tonic value. Tonic and melodic sequences were alternated, and there were 400 of each (200 per tonic). Within a given sequence, each note had a length of 650 ms and was ramped on and off with 5- and 10-ms Hanning windows, respectively. There were no pauses between adjacent notes within a sequence; the inter-sequence interval varied randomly between 1300 and 1350 ms. The total duration of the experiment was 55 min and the overall stimulation level was 70 dB SPL.

The neuromagnetic activity in response to the four note sequences was measured with a Neuromag-122 (Elekta Neuromag Oy) whole-head MEG system, at a 1000 Hz sampling rate, having been lowpass filtered 330 Hz. Ten normal-hearing listeners (mean age: 33.9 ± 12.01 years) participated in the study. Subjects watched a silent movie, while their AEFs were recorded; they were instructed to direct their attention to the movie. Spatio-temporal source analyses (Scherg 1990) were performed offline using the BESA 5.2 software package (BESA Software GmbH). Separate source models with one equivalent dipole in each hemisphere were created to analyze the P50m (using unfiltered data) and the sustained field (SF) (using data lowpass filtered at 1 Hz). Then, a model with two dipoles per hemisphere was built to localize the N1m/P2m complex associated with energy onset at the start of each sequence, and to localize changes between adjacent notes within a given sequence (pooled notes 2–4). A 1–30 Hz bandpass filter was applied to the data for these four-dipole fits. Afterwards, unfiltered source waveforms were derived for each model for each participant to compute grand-average source waveforms for both the tonic and melodic sequences. A principal component analysis was performed on the last few milliseconds of the epoch to compensate for drift. Statistical evaluation of specific component differences was accomplished in MATLAB using the multivariate statistic t_{sum} from the permutation test of Blair and Karniski (1993). Latencies and amplitudes of the additional data set described in the discussion section were evaluated statistically using a bootstrap technique to derive critical t -values (Efron and Tibshirani 1993).

3 Results

The data were fitted using a 4-dipole model with 2 dipoles in each hemisphere. One pair focused on the N1m temporal region of the first note in each bar to locate the “onset generator”; the other focused on the N1m temporal region of the three adapted melodic notes of each trial (notes 2–4 in the melody bar) to locate the “melody generator”—a technique developed for analysing melody processing in Patterson et al. (2015). The melody generator was located in the medial part of Heschl’s gyrus (HG); the onset generator about 10 mm more posterior in PT—a common location for the onset N1m. The source waves for the left and right hemispheres were very similar so the waves from the two hemispheres were averaged, in each case. The anterior-posterior difference between the average source locations was significant [left hemisphere: $t(9)=4.21$, $p=0.0023$; right hemisphere: $t(9)=4.02$, $p=0.003$].

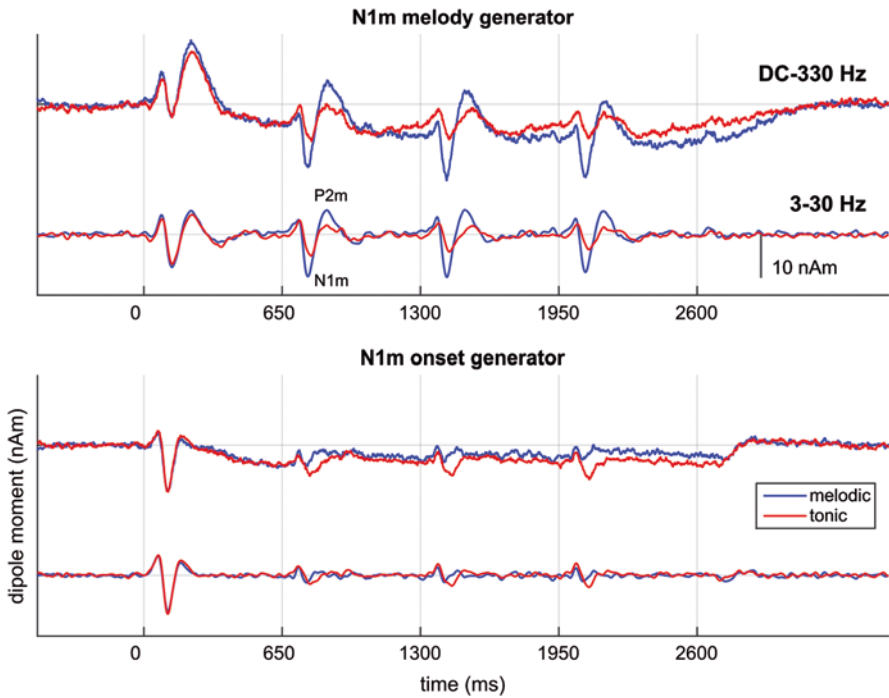


Fig. 1 Average source waves extracted from the data of the tonic (red) and melodic (blue) sequences using the melody generator (upper panel) and the onset generator (lower panel). The upper trace in each panel shows the unfiltered, grand average waveforms for the tonic (red) and melodic (blue) sequences. The sustained field is removed from the data in the lower trace of each panel by band pass filtering (3–30 Hz)

The upper panel of Fig. 1 shows the sources waves extracted with the melody generator separately for the tonic (red) and melodic sequences (blue) of each trial. The bottom panel shows the source waves extracted with the onset generator. The upper pair of curves are based on unfiltered data; the lower pair on data high-pass filtered at 3 Hz to remove the SF. The response to the first note of the tonic sequence is very similar to the first note of the melodic sequence. Thereafter, however, the adapted responses to notes 2–4 in the melodic sequence (blue) are substantially larger than their counterparts in the tonic sequence (red), both with regard to the N1m component and the P2m component (N100m: $t_{\max_note2} = -204$, $p = 0.002$; $t_{\max_note3} = -229.35$, $p = 0.002$; $t_{\max_note4} = -239.27$, $p = 0.002$; P200m: $t_{\max_note2} = 416.04$, $p = 0.0001$; $t_{\max_note3} = 427.91$, $p = 0.002$; $t_{\max_note4} = 443.30$, $p = 0.002$). The level, duration and timbre of the notes in the tonic and melodic sequences are the same, so the differences in the source waves reflect the processing of the sequence of pitch changes that together constitute the melody. The lower panel shows the source waves extracted with the posterior source (the onset generator). The response to the first note of the tonic sequence is once again very similar to the first note of the melodic sequence. The responses to succeeding notes are greatly attenuated,

illustrating the value of the 4-dipole analysis in the isolation of the energy-onset response.

In the melodic source generator (upper panel), the SF builds up (negatively) over about 500 ms following the P2m to the first note, from about 250 to 750 ms after the onset of the sequence. The magnitude of the SF decreases down to about -20 nAm below baseline for both the tonic and melodic sequences; it appears to get a little more negative towards the end of the melodic sequence, but not a great deal. Then, following the P2m of the last note, the SF decays back up to the baseline level over about 500 ms in both the tonic and melodic sequences. Statistically, over the course of the entire sequence, the SFs for the tonic and melodic conditions are not different, so it appears that the process that produces the SF does not distinguish the condition in which the notes produce a melody (SF differences based on melodic sources: -1973.99 , $p=0.29$; SF based on the 2-dipole SF-fit (0–1 Hz): $t_{\max} = 1236.12$, $p=0.545$).

The lower panel of Fig. 1 shows the sources waves extracted with the onset generator. In this posterior source, the responses to the initial note of the tonic and melodic sequences is virtually identical ($t_{\max} = 21.81$, $p=0.688$). Small differences are observed between the tonic and melodic responses to the second and succeeding notes in the unfiltered data. However, they are mainly due to a small difference in the SF which is not itself significant. When the SF is removed by filtering, the differences between the source waves are minimal. Thus, the SF in the posterior source does not differ for melodic sequences.

4 Discussion

The transients produced by the adapted notes (2–4) in the current experiment have consistently larger magnitudes in the melodic source generator. This prompted us to return to the melody experiment of Patterson et al. (2015) and do a more detailed analysis of the transient responses in that four-bar experiment, where there were two tonic bars and two melodic bars in each trial. The six adapted responses from the bars with the same sequence type (tonic or melodic) were shown to be very similar in that paper, and this made it possible to do a powerful, four-dipole fit with two dipoles focusing on the N1m temporal window and two focusing on the P2m temporal window. The fit isolated melodic and tonic generators in HG in both hemispheres, near the position of the melodic generator in the current experiment.

Figure 2 presents a comparison of the average, adapted responses in the data of Patterson et al. (2015) to the tonic notes from bars 1 and 3 (red lines) and the melodic notes of bars 2 and 4 (blue lines). The left and right panels show the TRs from the source waves obtained with the N1m and P2m temporal windows respectively. The figure shows that the technique reveals a double dissociation; the N1m is prominent in the N1m temporal window (left panel) but the P2m is not; whereas the P2m is prominent in the P2m temporal window (right panel), but N1m is not.

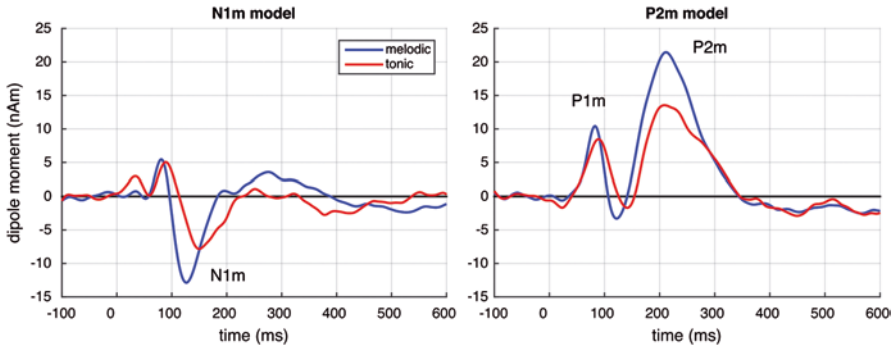


Fig. 2 Average transient responses from the 4-dipole model of Patterson et al. (2015) derived with an N1m temporal window (*left*) and a P2m temporal window (*right*)

Moreover, the adapted melodic notes (blue) produce a larger P2m (right panel) *and* a larger N1m (left panel) than the adapted tonic notes (red).

The P1m is larger with the P2m temporal window (right panel) but this is probably just because there is no N1m to reduce the apparent magnitude of the P1m with the P2m temporal window.

Both panels of Fig. 2 also show that the N1m peak occurs earlier in the *melodic* TR (blue), whereas one might have thought that it would take longer to produce a melodic response than a tonic response (latency diff_{passive} = 25.63 ms, crit- t = 14.15, $p < 0.05$; latency diff_{active} = 14.22 ms, crit- t = 8.21, $p < 0.05$). Although the peak delay differences might seem small, the timing of the N1m peak is one of the most precise statistics in MEG data, and these differences are highly significant—more so than the N1m amplitude differences in these panels (amplitude diff_{passive} = 2.74 nAm, crit- t = 2.6, n.s.; amplitude diff_{active} = 5.52 nAm, crit- t = 5.35, $p < 0.05$).

5 Conclusions

The shorter melody experiment presented at the start of this paper shows how the classic energy-onset generator in PT can be separated from the melody generator in the anterior part of HG. The transients obtained from the melody generator contain both N1m and P2m components, and both components are larger in melodic sequences than they are in tonic sequences. Moreover, the N1m peak delay is shorter in melodic sequences. Reanalysis of the TR data from the four-bar melody experiment suggests that the N1m and P2m components of the response in HG can be separated with a 4-dipole source model, and that melody processing is more readily observed in the shape and magnitude of the P2m response. Finally, the sustained field originates from a generator in HG not far from the melody generator; however, the SF itself does not distinguish between the tonic and melodic sequences.

Open Access This chapter is distributed under the terms of the Creative Commons Attribution-Noncommercial 2.5 License (<http://creativecommons.org/licenses/by-nc/2.5/>) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

The images or other third party material in this chapter are included in the work's Creative Commons license, unless indicated otherwise in the credit line; if such material is not included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt or reproduce the material.

References

- Blair RC, Karniski W (1993) An alternative method for significance testing of waveform difference potentials. *Psychophysiology* 30:518–524
- Efron B, Tibshirani RJ (1993) *An introduction into the bootstrap*. Chapman & Hall, New York
- Goto M, Hashiguchi H, Nishimura T, Oka R (2003). RWC music database: music genre database and musical instrument sound database. In: Hoos HH, Bainbridge D (eds) *Proceedings of the 4th International Conference on Music Information Retrieval*, pp 229–230
- Kawahara H, Irino T (2004) Underlying principles of a high-quality speech manipulation system STRAIGHT and its application to speech segregation. In: Divenyi PL (ed) *Speech separation by humans and machines*. Kluwer Academic, Dordrecht, pp 167–180
- Kawahara H, Masuda-Kasuse I, de Cheveigne A (1999) Restructuring speech representations using pitch-adaptive time-frequency smoothing and instantaneous-frequency based F0 extraction: possible role of repetitive structure in sounds. *Speech Comm* 27:187–204
- Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B (2003) Neuro-magnetic evidence for a pitch processing centre in Heschl's gyrus. *Cereb Cortex* 13:765–772
- Patterson RD, Uppenkamp S, Johnsruide I, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. *Neuron* 36:767–776
- Patterson RD, Uppenkamp S, Andermann M, Rupp A (2015) Brain imaging the activity associated with pitch intervals in a melody. *Proc Meet Acoust* 21:050009
- Rupp A, Uppenkamp S (2005). Neuromagnetic representation of short melodies in the auditory cortex. In: Fastl H, Fruhmann M (eds) *Fortschritte der Akustik-DAGA 2005*. DEGA e.V., Berlin, pp 473–474
- Scherg M (1990) Fundamentals of dipole source potential analysis. In: Grandori F, Hoke M, Romani GL (eds) *Auditory evoked magnetic fields and potentials*. *Advances in audiology* 5. Karger, Basel, pp 40–69
- Uppenkamp S, Rupp A (2005). Functional MR imaging of the processing of pitch changes in human listeners. In: Fastl H, Fruhmann M (eds) *Fortschritte der Akustik-DAGA 2005*. Deutsche Gesellschaft für Akustik e.V., Berlin, pp 471–472