

A neurophysiological study into the foundations of tonal harmony

Elika Bergelson^{a,c} and William J. Idsardi^{a,b}

Our findings provide magnetoencephalographic evidence that the mismatch-negativity response to two-note chords (dyads) is modulated by a combination of abstract cognitive differences and lower-level differences in the auditory signal. Participants were presented with series of simple-ratio sinusoidal dyads (perfect fourths and perfect fifths) in which the difference between the standard and deviant dyad exhibited an interval change, a shift in pitch space, or both. In addition, the standard-deviant pair of dyads either shared one note or both notes were changed. Only the condition that featured both abstract changes (interval change and pitch-space shift) and two novel notes showed a significantly larger magnetoencephalographic mismatch-negativity response than the other conditions in the right hemisphere. Implications for music and language processing are discussed. *NeuroReport*

20:239–244 © 2009 Wolters Kluwer Health | Lippincott Williams & Wilkins.

NeuroReport 2009, 20:239–244

Keywords: auditory processing, dyads, intervals, magnetoencephalography, mismatch negativity, music, tone, vowel formants

^aDepartment of Linguistics, ^bNeuroscience and Cognitive Science Program, University of Maryland, Maryland and ^cDepartment of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence to Elika Bergelson, Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut, Suite 400A, Philadelphia, PA 19104-6228, USA

Tel: +1 614 598 6937; fax: +1 215 573 9247; e-mail: elika.bergelson@gmail.com

Received 28 September 2008 accepted 8 October 2008

Introduction

Although there is a wealth of behavioral literature on human processing of simple and complex musical events [1,2], and an emerging literature on the neurophysiological responses to auditory stimuli (Jenkins J, Poeppel D, Idsardi W, in preparation; [3–5]), we have yet to provide an adequate account of how the brain processes the building blocks of complex coherent auditory events such as musical chords.

Earlier behavioral research [6] has shown that both adults and infants are better at discriminating changes to a series of intervals when the intervals used have simple frequency ratios, such as perfect fourths and perfect fifths. Other research has shown a role for directionality in analysis of mistuned intervals [7]. There has also been much research on three-note chords and chord progressions showing that listeners have very specific expectations about which chords are licit given a certain key and chordal context [1,2] (for a comprehensive summary, see Ref. [8]). In comparison, on the neurophysiological side, although many researchers have looked at the brain responses to single tones, phonemes, or clicks [9–13], as well as words, complex chords, or phrases [14,15], thus far there is a relative dearth of studies examining basic auditory responses to the simplest sort of simultaneous musical expressions, two-tone stimuli (Jenkins J, Poeppel D, Idsardi W, in preparation).

In this study, we examine two-note chords or dyads. Dyads serve as a particularly good starting point for

understanding complex auditory signals, as they are the smallest harmonically complex musical unit. Moreover, examining how our brains process dyads can help improve our understanding of vowel processing, given that two formants are the minimum required for vowel identification [16]. In this study, we investigate the difference in electrophysiological response to dyads with simple ratios in a mismatch-negativity paradigm [17]. We examined an early, automatic, and robust response to oddball stimuli: the magnetoencephalographic mismatch-negativity field (MMNm). The MMNm functions as a preattentive difference detector, and is activated by deviants in a stream of standards that can vary over any number of dimensions in various modalities (for an overview, see Ref. [18]). The present investigation uses the MMNm response as a tool to measure how, when, and by how much our electrophysiological response is modulated by various abstract and concrete differences between a dyad when it serves as a standard versus when it serves as a deviant. Although some earlier research has looked at the MMNm response to musical intervals [19], none has yet looked at the role of interval or pitch shift in relation to the size of the MMNm.

Given the robust nature of the MMNm, we were able to parametrically vary our dyad pairs on two dimensions: the size of the interval (either a perfect fourth or a perfect fifth, with 4:3 and 3:2 frequency ratios, respectively) and its location in pitch space. As a result of the construction of the Western musical system, this meant that both the pitch-space shifts and one of the combined

pitch-space and interval-change conditions introduced two novel notes in the second dyad, whereas the interval change and one of the pitch-space and interval-change conditions only introduced one novel note in the second dyad.

We predicted that every condition would elicit a mismatch response. In addition, two patterns of responses are initially plausible: one consistent with a signal-processing account and another relying on abstract cognitive categories. The signal-processing account would predict the MMNm would be larger in the three conditions in which the deviant dyad stimulus contained two novel notes rather than one (CF-DA, CF-DG, and CG-DA). In contrast, the abstract category account would predict a larger MMNm response when the deviant dyad changed both the interval and the pitch space, (CF-DA and CG-DG).

Materials and methods

Participants

Thirty-two (18 female; age range=19–48 years; mean age=27.5 years) adult volunteers participated in this study. Most ($n=30$) tested were strongly right handed on the Edinburgh Handedness Survey (scoring greater than 8/10) the rest ($n=2$) tested equally right and left handed (these two volunteers were tested in conditions CG-DA and DA-DG, see below) [20]. All participants gave written informed consent, and either volunteered their time ($n=22$), or received course credit ($n=10$). Each session lasted for 60–90 min. The involvement of human participants in the reported experiment was approved by the Institutional Review Board of the University of Maryland (College Park, Maryland, USA).

Stimuli

Auditory stimuli were pure sinusoids, synthesized with Praat [21] at a sampling frequency of 44.1 kHz. Each stimulus featured two simultaneous pitches from the conventional Western music scale and lasted for 250 ms with the amplitudes of the stimuli normalized to 75 dB. As the design was testing the MMNm elicited by different pairs of dyads, we systematically varied the ratio of the frequencies and their location in pitch space. The frequency ratios of the dyads were either those of a perfect fourth, 4:3, or perfect fifth, 3:2. These ratios were picked because of their easy discriminability in behavioral work and prevalence in music of the world [5,22]. Furthermore, our choice of using perfect fourths and perfect fifths in our study is in line with earlier behavioral findings of these intervals' easy discriminability. The pairs of notes were CF, 262 and 349 Hz; CG, 262 and 392 Hz; DG, 294 and 392 Hz; and DA, 294 and 440 Hz, and are summarized in Table 1; the relation between conditions is shown schematically in Figure 1.

Procedure

Participants were made to lay in supine position in a magnetically shielded room while magnetoencephalographic recordings were made using a 157-channel whole-head axial gradiometer MEG system (Kanazawa Institute of Technology, Kanazawa, Japan). Etymotic ER3A insert earphones (Etymotic Research Inc., Elk Grove Village, Illinois, USA) delivered the auditory stimuli binaurally; these earphones were calibrated to have a flat frequency response between 50 and 3100 Hz within the shielded room. Auditory stimuli were presented with PsyScope [23] in four blocks of 7.5 min each, with optional breaks between each block. Each participant heard one pair of dyads in an oddball paradigm, a series of standards followed by a deviant. Blocks 1a and b featured the dyads in one order, whereas blocks 2a and b flipped this order. The ISI and the number of standards in a row varied randomly (500–1000 ms, and 4–7, respectively). Participants heard 780 standards and 120 deviants in each dyad order. Participants laid passively and could elect to watch a video (with no audio). Each participant heard only one of the dyad pairs (e.g. CF-DA), thus making this experiment a fully between-subjects design.

The signal was sampled at DC with an online 200 Hz low-pass filter and 60 Hz notch filter. Data were noise-reduced and band-pass filtered offline. The noise reduction used a multishift PCA noise reduction algorithm [24], and the band-pass filtration occurred with a Hamming-window digital filter with 0.5 and 30 Hz frequency cut-offs.

Before the experiment, volunteers each participated in a two-tone pretest, in which they heard a series of approximately 200 sinusoidal tones, 100 tones at 1000 and 250 Hz each, presented in random order. This was done to ensure that each participant had a strong M100 response and a standard M100 field contour, indicating that the response is presumably located in auditory cortex.

The five source and sink channels in each hemisphere that provided the strongest M100 response to these pure sinusoids were selected from the magnetic contour map. A total of 20 analysis channels were selected for further analysis based on the magnitude of the M100 response to these sinusoidal tones. For each participant, the sinusoid generating the strongest overall M100 response was used, and the strongest five source and sink channels in each hemisphere for that tone were selected for use in the analysis of the test data. Therefore, all channels for analysis for each participant were selected on the basis of the pretest tones, before the analysis of the experimental data. The root mean square average of these channels was used for all further analyses of the MMNm. We measured the latency and amplitude of the averaged deviant responses, and the area of difference between the response to each dyad when it was functioning as the standard and as the deviant.

Table 1 Summary of conditions

Dyad pair	Interval	Number of novel notes	Type of change
CF–CG	P4–P5	1	Interval
CF–DG	P4–P4	2	Pitch space
CF–DA	P4–P5	2	Interval and pitch space
DG–DA	P4–P5	1	Interval
CG–DA	P5–P5	2	Pitch space
CG–DG	P5–P4	1	Interval and pitch space

P4, perfect fourth; P5, perfect fifth.

Fig. 1

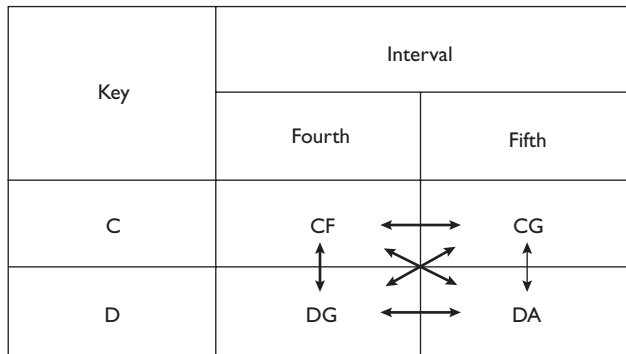


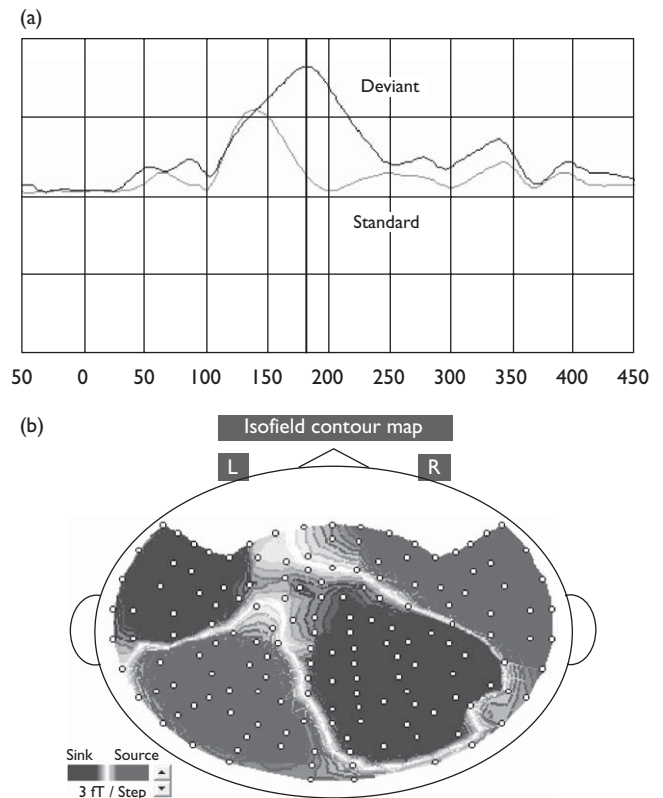
Diagram of conditions. Each arrow connecting a pair of dyads represents one tested condition.

Results

As expected, each condition elicited an MMNm. Figure 2 shows a representative response to a given dyad as standard and deviant. The distributions of the responses were examined for approximate normality and 13 Tukey outliers were excluded (points beyond the first or third quartile by more than 1.5 times the interquartile range). To compare responses across hemispheres, matched-pair *t*-tests within participants were calculated. Overall, the left hemisphere exhibited stronger [$t(55)=2.90, P < 0.005$] but slower [$t(55)=4.65, P < 0.0001$] deviant responses, and a larger difference response [$t(58)=5.84, P < 0.0001$].

To analyze the responses to the various experimental conditions, the left and right hemispheres were analyzed separately for the three dependent measures (deviant amplitude, deviant latency, and amplitude difference) with dyad type condition and mismatch presentation direction as factors, for a total of six mixed-effects analyses of variance (ANOVAs) of the form response – condition × direction, with participant as a random effect. We found no main effects for presentation direction across all measures (all $P > 0.34$), as well as no interaction of condition and direction (all $P > 0.60$). Thus, given that there was no asymmetry of mismatch field response in any condition (i.e. whether the dyads were presented in an order such as CF–DA or DA–CF was irrelevant),

Fig. 2



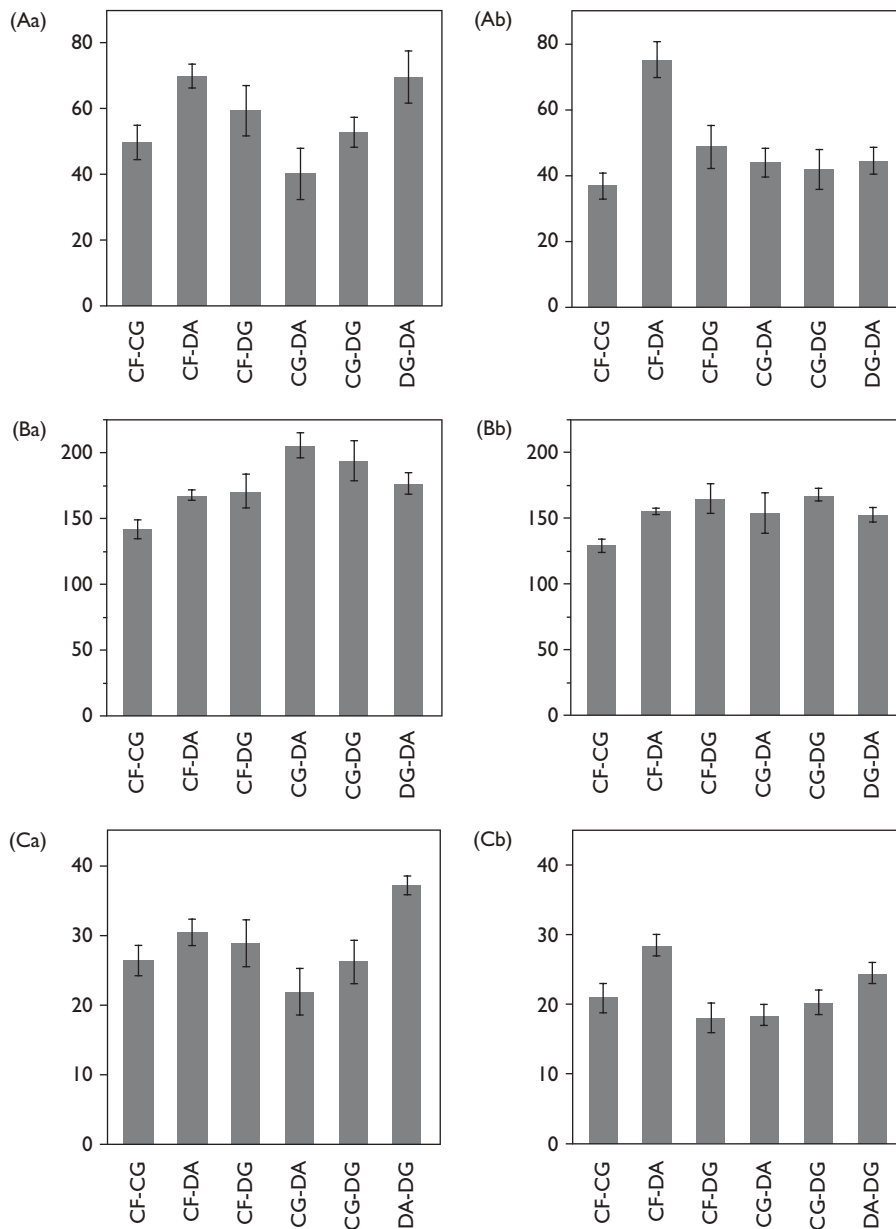
Magnetoencephalographic analysis. (a) A representative participant's response to a dyad as deviant (dark line) versus standard (light line); (b) scalp distribution at peak magnetic mismatch field. The contour plot shows that the underlying evoked field is likely generated over auditory cortex, as this pattern matches the pattern generated by the pretest M100 response to a single tone.

responses were collapsed across direction for further analyses using simple one-way ANOVAs of the form response – condition. The ANOVAs revealed that all measures (deviant amplitude, deviant latency, and amplitude difference) in both hemispheres showed significant main effects for the pairs tested [all $F(5,50) > 3.12, all P < 0.016$], but showed different patterns of significant differences in post-hoc pairwise comparisons [corrected for multiple comparison using the Tukey Honestly Significant Differences (HSD) procedure]. Means and standard errors for peak deviant amplitude, latency, and difference measures in both hemispheres are shown in Fig. 3.

Peak deviant amplitude

In the right hemisphere, the CF–DA condition generated the largest peak deviant MMNm response amplitude, which was significantly stronger than all other conditions (Tukey HSD correction for multiple comparisons, $P < 0.05$). In the left hemisphere, CF–DA also had the highest amplitude, but this was statistically significant only

Fig. 3



(A) Mean deviant peak amplitude (fT) for (a) left hemisphere, (b) right hemisphere. (B) Mean deviant peak latency (ms) for (a) left hemisphere, (b) right hemisphere. (C) Root mean square difference amplitude (fT) between 100 and 300 ms for (a) left hemisphere, (b) right hemisphere. In all charts the error bars indicate one standard error of the mean.

when compared with CG-DA (Tukey HSD correction for multiple comparisons, $P < 0.05$).

Peak deviant latency

The CF-CG condition yielded marginally faster peak deviant latencies than the other conditions in both hemispheres, but this did not reach statistical significance. Otherwise, there was no clear pattern in the ordering of the latencies of the conditions between the two hemispheres.

Difference

In the right hemisphere, the CF-DA condition generated the largest difference response (deviant - standard), which was significantly stronger than three other conditions (CG-DG, CG-DA, and CF-DG, Tukey HSD correction for multiple comparisons, $P < 0.05$). In the left hemisphere, DG-DA had the largest difference response, which was significantly stronger than three other conditions (CG-DG, CG-DA, and CF-CG), but

was not significantly larger than the second largest response, that for CF–DA.

In summary, one condition stands out, CF–DA. In the right hemisphere, the amplitude of the deviant response to CF–DA is clearly stronger than all other conditions, and CF–DA is the strongest or second-strongest response in all other deviant amplitude and difference measures. Moreover, the CF–DA condition is notable, in that it is the only condition that includes an interval change, a shift in pitch space, and two novel notes.

Discussion

Our results show that by 150 ms the brain is able to track the magnitude of changes in dyads. Moreover, these results show that the MMNm does not track changes solely on the basis of the acoustic novelty of the deviant auditory signal, nor solely on the basis of an abstract class difference. Rather, only in the condition that included the largest acoustic novelty (two new notes) and both kinds of abstract changes (an interval change and a shift in pitch space) did the right hemisphere show a significantly greater MMNm. These results are in line with the musical-processing literature showing that fine pitch discrimination occurs in the right hemisphere [5].

In addition, these results show an interesting parallel with the speech perception literature, and language more broadly construed. Just as is found in vowel perception, we have shown that the MMNm is sensitive and able to meld various frequencies and determine differences across tokens robustly and quickly using pure tones as well. This can be taken to show just how rapidly our auditory cortex computes properties of incoming auditory stimuli not only when they matter for meaning and language comprehension (as when we tell vowels apart), but also when meaning or a labeling system is absent (as in dyads). This extends the findings of others in the music, pitch, and speech-tone literature [25–30]. The nature of the brain response to sine waves as opposed to other timbres is also of great interest [31]. Moreover, it seems that the way pitch is processed depends in part on the context involved [32]; this is a line of research we hope to extend with currently ongoing work.

Furthermore, our results show the MMNm to be sensitive to both abstract and acoustic differences; the same multilevel sensitivity is present in language. The MMNm reflects low-level linguistic differences, such as vowel or phoneme changes [16,17], but is also activated for words among pseudowords [15]. Thus, our study also hints at evidence for parallels in how the brain treats music and language on multiple levels simultaneously, perhaps suggesting that all auditory input undergoes similar computational operations.

Conclusion

This study has improved our understanding of how the brain processes dyads. We found that the greatest brain response was elicited in the condition that crucially had three features: (i) a key change, (ii) an interval change, and (iii) two novel notes. This shows the MMNm to be more than an unsophisticated monitor of acoustic change, yet less than a clear marker of abstract categories. This research will thus serve as a foundation from which to further examine how the brain analyzes incoming linguistic and musical signals, that is, auditory signals consisting of two or more simultaneous frequencies.

Acknowledgements

The authors thank Jeff Walker for his tremendous help in all practical aspects of running this project, as well as David Poeppel, and Phil Monahan for comments and suggestions and Kevin de Souza for initial discussions. This research was supported by NIH R01 05660, awarded to David Poeppel (project director) and William Idsardi (coprincipal investigator).

References

- 1 Krumhansl CL. *Cognitive foundation of musical pitch*. Oxford: Oxford University Press; 1990.
- 2 Krumhansl CL. The cognition of tonality – as we know it today. *J New Music Res* 2005; **33**:253–268.
- 3 Sams M, Paavilainen P, Alho K, Naatanen R. Auditory frequency discrimination and event-related potentials. *Electroencephalogr Clin Neurophysiol* 1985; **62**:437–448.
- 4 Tervaniemi M, Kujala A, Alho K, Virtanen J, Ilmoniemi RJ, Näätänen R. Functional specialization of the human auditory cortex in processing phonetic and musical sounds: a magnetoencephalographic (MEG) study. *Neuroimage* 1999; **9**:330–336.
- 5 Zatorre RJ, Belin P, Penhune V. Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 2002; **6**:37–46.
- 6 Schellenberg EG, Trehub SE. Natural intervals in music: a perspective from infant listeners. *Psychol Sci* 1996; **7**:272–277.
- 7 Schellenberg EG. Asymmetries in the discrimination of musical intervals: going out-of-tune is more noticeable than going in-tune. *Music Percept* 2001; **19**:223–248.
- 8 Patel AD. *Music, language, and the brain*. Oxford: Oxford University Press; 2008.
- 9 Roberts TPL, Poeppel D. Latency of auditory evoked M100 as a function of tone frequency. *Neuroreport* 1996; **7**:1138–1140.
- 10 Gage N, Poeppel D, Roberts TPL, Hickok G. Auditory evoked M100 reflects onset dynamics of speech sounds. *Brain Res* 1996; **814**:236–239.
- 11 Kazanina N, Phillips C, Idsardi W. The influence of meaning on the perception of speech sounds. *Proc Natl Acad Sci U S A* 2006; **103**:11381–11386.
- 12 Ohl FW, Scheich H. Orderly cortical representation of vowels based on formant interaction. *Proc Natl Acad Sci U S A* 1997; **94**:9440–9444.
- 13 Brugge JF, Volkov IO, Oya H, Kawasaki H, Reale RA, Fenoy A, et al. Functional localization of auditory cortical fields of human: click-train stimulation. *Hear Res* 2008; **238**:12–24.
- 14 Leino S, Brattico E, Tervaniemi M, Vuust P. Representation of harmony rules in the human brain: further evidence from event-related potentials. *Brain Res* 2007; **1142**:169–177.
- 15 Shtyrov Y, Pulvermüller F. Neurophysiological evidence of memory traces for words in the human brain. *Neuroreport* 2002; **13**:521–525.
- 16 Delattre PC, Liberman AM, Cooper FS, Gerstman LJ. An experimental study of the acoustic determinants of vowel color: observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word* 1952; **8**:195–210.
- 17 Näätänen R. *Attention and brain function*. Hillsdale, New Jersey: Erlbaum; 1992.

- 18 Näätänen R, Paavilainen P, Rinne T, Alho K. The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 2007; **118**:2544–2590.
- 19 Trainor LJ, McDonald KL, Alain C. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J Cogn Neurosci* 2002; **14**:430–442.
- 20 Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971; **9**:97–113.
- 21 Boersma P, Weenink D. Praat: doing phonetics by computer (version 5.0.27) 2008; [Computer program].
- 22 Dowling WJ, Harwood DL. *Music cognition*. Orlando, Florida: Academic Press; 1986.
- 23 Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav Res Methods Instrum Comput* 1993; **25**:257–271.
- 24 Cheveigné AD, Simon JZ. Denoising based on time-shift PCA. *J Neurosci Methods* 2007; **165**:297–305.
- 25 Wu C, Kirk IJ, Hamm JP, Lim VK. The neural networks involved in pitch labeling of absolute pitch musicians. *Neuroreport* 2008; **19**:851–854.
- 26 Kaiser J, Heidegger T, Wibrall M, Altmann CF, Lutzenberger W. Alpha synchronization during auditory spatial short-term memory. *Neuroreport* 2007; **18**:1129–1132.
- 27 Mizuno T, Sugishita M. Neural correlates underlying perception of tonality-related emotional contents. *Neuroreport* 2007; **18**:1651–1655.
- 28 Otsuka A, Tamaki Y, Kuriki S. Neuromagnetic responses in silence after musical chord sequences. *Neuroreport* 2008; **19**:1637–1641.
- 29 Soeta Y, Nakagawa S. The effects of pitch and pitch strength on an auditory-evoked N1m. *Neuroreport* 2008; **19**:783–787.
- 30 Swaminathan J, Krishnan A, Gandour JT. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport* 2008; **19**:1163–1167.
- 31 Meyer M, Baumann S, Jancke L. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage* 2006; **32**:1510–1523.
- 32 Otsuka A, Kuriki S, Murata N, Hasegawa T. Neuromagnetic responses to chords are modified by preceding musical scale. *Neurosci Res* 2008; **60**:50–55.