Experimental Brain Research (2019) 237:2811–2828
https://doi.org/10.1007/s00221-019-05638-4

RESEARCH ARTICLE

Event-related potentials to single-cycle binaural beats of a pure tone, a click train, and a noise

Pekcan Ungan[1] · **Suha Yagcioglu**[2] · **Ece Ayik**[3]

Received: 6 April 2019 / Accepted: 19 August 2019 / Published online: 26 August 2019

Abstract

There are only few electrophysiological studies on a phenomenon called "binaural beats" (BBs), which is experienced when two tones with frequencies close to each other are dichotically presented to the ears. And, there is no study in which the electrical responses of the brain to BBs of complex sounds are recorded and analyzed. Owing to a recent method based on single-cycle BB stimulation with sub-threshold temporary monaural frequency shifts, we could record the event-related potentials (ERPs) to BBs of a 250-Hz tone as well as those to the BBs of a 250/s click train and to the BBs of a recurrent 4-ms Gaussian noise. Although fundamental components of the click train and noise stimuli were lower in intensity than the tonal stimuli in our experiments, the N1 responses to the BBs of the former two wide-spectrum sounds were recorded with significantly larger amplitudes and shorter latencies than those to the BBs of a tone, suggesting an across-frequency integration of directional information. During a BB cycle of a complex sound, the interaural time differences (ITDs) of the spectral components are all equal to each other at any time; whereas their interaural phase differences (IPDs) are all different. The ITD rather than the IPD should, therefore, be the cue that is relied upon by the binaural mechanism coding the perceived lateral shifts of the sound caused by BBs. This is in line with across-frequency models of human auditory lateralization based on a common ITD, fulfilling a straightness criterion.

Keywords Binaural beats · Click train · Noise · ERP-N1 · Sound lateralization · Across-frequency integration

Introduction

When two tones with frequencies close to each other are delivered dichotically to the ears, the listener perceives a sort of modulation in the sound. This phenomenon, which is known for more than a century (Rayleigh 1907; Stewart 1917) and called "binaural beats" (BBs), results from binaural interaction of the cochlear neural codes in the brain. Besides several non-invasive brain stimulation methods such as transcranial electric stimulation or pulsed transcranial ultrasound stimulation, BB stimulation is becoming an alternative technique that could be used to induce EEG phase synchronization in the brain and to study or modify cognitive brain function (Gao et al. 2014; Ioannou et al. 2015; Becher et al. 2015; Beauchene et al. 2016; Hommel et al. 2016; Chaieb et al. 2017; Garcia-Argibay et al. 2019; Derner et al. 2018). Further basic studies are, therefore, necessary to understand the neurophysiological aspects of this phenomenon better and to search for the optimal conditions that could improve its effectiveness as a non-invasive stimulus modulating the brain's electrical activity (EEG).

Any auditory event elicits in the brain some electrical potentials which can be recorded non-invasively from the scalp. These potentials, which are time- or phase-locked to the auditory event causing them, may be steady state or transient depending on whether the stimulus is a steady-state or a transient one. When the brain is stimulated by a transient auditory event, around the 100–200 ms time period following that event, an EEG response called N1–P2 complex is

✉ Pekcan Ungan
pekungan@gmail.com

Suha Yagcioglu
yagcioglu.s@gmail.com

Ece Ayik
eceayik@gmail.com

1 Department of Biophysics, School of Medicine, Koc University, Istanbul, Turkey

2 Department of Biophysics, Faculty of Medicine, Hacettepe University, Ankara, Turkey

3 Graduate School of Science and Engineering, Koc University, Istanbul, Turkey

recorded from the human scalp. The vertex-negative N1 deflection of this response complex, which is recorded with maximal amplitudes at the fronto-central scalp electrodes (Vaughan and Ritter 1970; Picton et al. 1974), is the most studied and typical auditory event-related potential (ERP) component and has been shown to reflect stimulus features such as its intensity, spatial origin, duration, and timing (Näätänen and Picton 1987). It has been shown by numerous EEG/MEG and lesion studies that the cortical source of the auditory N1 response is on planum temporalis superior (e.g., Papanicolaou et al. 1990; Scherg et al. 1989; Knight, et al. 1988). The EEG responses evoked by BBs, which is also an auditory event, have been shown to be cortical in origin (Pratt et al. 2010), although early stages of the binaural interaction underlying this auditory event take place in the brainstem (Wernick and Starr 1968; Ungan and Yagcioglu 2002).

However, conventional BBs are steady state and recurrent with relatively high repetition rates, in general. And, when they are used as an auditory stimulus, the responses which they elicit in the auditory cortex can be recorded with very small amplitudes hardly reaching 1 μV (e.g., Pratt et al. 2009), because these frequent stimuli do not allow the cortical responses to recover adequately. When the repetition rate is low, this time, the rate of change in interaural phase remains too slow to elicit a notable auditory-evoked response because of its well-known rate sensitivity. Furthermore, in the steady-state EEG responses (SSRs) evoked by conventional BB stimuli, the responses to each of the recurrent beats intermingle, making their identification, physiological interpretation, and source analysis very difficult. It is for this reason that identification and measurement of the N1–P2 complex, which has been repeatedly shown to be a typical long-latency response component originated in the auditory cortex (e.g., Scherg et al. 1989; Papanicolaou et al. 1990; Ungan et al. 2001), are not possible with steady-state BB responses.

Using a method designed and tested in our laboratory, it has been possible to overcome the above-mentioned issues and to record reliably the transient responses of the cortical mechanisms processing BBs (Ungan et al. 2019). This method, which was based on the single-cycle BB stimulation suggested by Özdamar et al. (2011), enabled us to record, identify, and analyze the auditory cortical ERP components with different latencies, which would have intermingled when recorded by using steady-state BB stimulation. Owing to the single-cycle BB stimulation technique, inter-beat interval could be increased significantly without slowing the ITD change during the beats, and the cortical responses could be recorded with amplitudes several times larger than it would be possible with steady-state BB stimulation using comparable cohorts of subjects and stimulus parameters. In Pratt et al. 2009, for example, the peak-to-peak amplitude of the grand average ($N = 18$; 18–24 years of age) responses to 3-Hz steady-state BBs of a 250-Hz base tone with 70 dB nHL is around 1 μV, which is about four times smaller than the grand average amplitude of the single-cycle BB responses reported by Ungan et al. (2019). By comparing the amplitude ratios between the N1-P2 response and remnant pre-stimulus EEG of the grand average ERPs presented in these two studies, it was also shown in the latter study that the responses evoked by single-cycle BB had a higher SNR than those evoked by steady-state BBs with similar stimulus parameters regarding the interaural frequency difference and the frequency and intensity of the base tone.

As was also mentioned in our previous study (Ungan et al. 2019), an important issue encountered in recording the ERPs to single-cycle BBs is that the temporary shifts that are made in the frequency of the sounds in the two ears may also excite the monaural mechanisms that are sensitive to frequency modulation (FM). In such cases, it may be necessary to eliminate this monaural FM response superimposed on the BB response, by recording in a control experiment the monaural FM response alone and subtracting this FM response from the compound response, as was attempted in the work of Özdamar et al. (2011). However, if the amount of monaural frequency shifts for BBs is kept below the FM perception threshold, as we did in our previous study mentioned above, it becomes possible to avoid the confounding effect of monaural FM responses which would have been elicited otherwise. Using the same method in the present study, we were able to record the specific BB responses directly; i.e., with no need for additional recording of the monaural FM responses and subtracting them from the recorded ERP.

The percept of BB was originally described for pure tones and it is associated with a neural mechanism based on the phase differences between the tones in the two ears (i.e., interaural phase difference, IPD) (Pratt et al. 2010). Indeed, when steady tones with very small interaural frequency differences are used to create BBs, the phenomenon is perceived as a periodic movement of an intracranial sound from one ear to the other rather than a modulation in some feature of the sound (Perrott and Musicant 1977). However, there is still not a solid consensus as to whether IPD or ITD correlates better with the human perception of sidedness, although it has been known for more than a century that lateralization of low-frequency tones is mainly based on interaural phase or time differences. There are findings, suggesting that IPD may be more fundamental physiologically (McAlpine et al. 2001). It was reported by Yost (1981) that the relationship between the lateral position of sine tones and IPD was independent of frequency, whereas the perceived lateralization for an interaural time difference (ITD) varied depending on frequency of the tone. In contrast with similar reports suggesting that the human binaural system is acting more as an IPD meter than as an ITD meter, there are studies showing that the azimuth of free-field sources correlates better with ITD than IPD (Kuhn 1979). In addition, there are across-frequency models for human lateralization of broad-band stimuli which are based on a common ITD in different frequency channels (Stern et al. 1988; Shackletona et al. 1992). Because IPD is simply equal to $2\pi f$ times the ITD, where "f" is the sound frequency, there is a direct relationship between the two for a pure tone. Therefore, a pure tone is not a convenient stimulus to compare the weights of IPD and ITD in sound lateralization. Actually, natural sounds and thus the auditory stimuli that we encounter in normal life are not pure tones but complex sounds having many spectral components. When the source of a complex sound is laterally shifted, the ITDs associated with all of its spectral components change equally, whereas the changes in the IPDs associated with them are totally different from each other. In other words, for a complex sound, a certain lateral position of the sound corresponds to different IPD values for different spectral components, but it corresponds to a single ITD value for all of them. It is important, therefore, to study the cortical responses elicited by BBs of complex sounds, which are created by gradual ITD shifts of their waveforms as a whole, and to see how these responses compare with those to BBs of a pure tone. Measuring stronger cortical responses to BBs of a complex sound than to BBs of a pure tone should indicate an across-frequency integration over ITD, supporting the hypothesis that the human sound lateralization mechanism uses the ITD cue rather than the IPD cue.

However, the BB studies in the literature using wide-band noise or multi-tone sounds are all behavioral (McFadden and Pasanen 1975; Saberi 1995; Bernstein and Trahiotis 1996; Akeroyd 2010). In connection with the theme of the present study, it is interesting to note that a strong percept of BB also with such complex sounds is reported in these studies. Akeroyd (2010), for instance, reported that there was a clear percept of motion for BBs of noise, especially when the noise is wide-band, and the percept was far stronger for a noise-BB than for a tonal one. However, we have not encountered any work in the literature in which the brain's electrical responses to BBs of complex sounds are studied. In fact, the brain's responses to a change in the interaural time relationship have been recorded and analyzed for complex sounds such as click trains and noise (Halliday and Callaway 1978; von Wedel 1982; Ungan et al. 1989; McEvoy et al. 1990; Jones et al. 1991: Sams et al. 1993; Chait et al. 2005; Edmonds and Krumbholz 2014), but in all these studies, the auditory event eliciting the responses were abrupt alterations in ITD or interaural coherence, rather than a gradual change in those interaural parameters. Although largely overlapping binaural mechanisms of the brain may be involved in the processing of the transient changes in interaural time relationship used in these studies and the present one, it seems that these stimuli are somewhat different from each other perceptually. According to the behavioral BB studies of Perrott and Musicant (1977) and Grantham and Wightman (1978), for a clear moving sound sensation, the beat cycle length would be at least 1 s. In line with these findings, in a recent study of our group (Ungan et al. 2019), the percept created by single-cycle BBs with a cycle length as short as 322 ms or by each beat of a steady-state BB repeating with such short periods is not of a laterally moving sound, but a vague auditory event resembling to a change in loudness. An abrupt change in ITD, on the other hand, is known to be perceived as a shift in the lateral position of the intracranial sound image, even when the ITD change lasts for a period as short as 50 ms (e.g., Ungan et al. 2001). Consequently, the neural mechanisms processing the BB stimulus that we use and those processing an abrupt ITD change used in the above-cited studies may not exactly be the same, although the stimuli in all these cases are basically an ITD transition. Therefore, to obtain and analyze the cortical responses to BBs of complex sounds for the first time, we have recorded in the present study the ERPs to BBs of click trains and Gaussian noise beside those to BBs of a 250-Hz sine tone.

Actually, the present BB-ERP study is a continuation of the earlier studies by Ungan et al. (1989, 2001) on the ERPs to changes in the ITD of complex sounds. In the former one of these studies, the laterality-reversal ERPs were recorded in response to abrupt alteration of the ITD of dichotically delivered 70-Hz trains of 1-ms clicks, and the robust N1-P2 responses elicited were described. In the latter study, the ERPs to 50-ms lateral shifts of a complex sound image were recorded using a novel stimulus in the form of square wave modulated 100-Hz trains of 1-ms clicks. Lateral shifts were created by introducing an ITD or an ILD (interaural level difference), and the ERPs to these transient ITD and ILD stimuli were recorded. The N1 components of these ERPs were compared with each other regarding their component structures, scalp topographies, and cortical dipolar sources. The change in interaural disparities was abrupt in these studies. In the present study, on the other hand, BBs created by a gradual change in ITD are used as stimuli, and the ERPs to BBs of complex sounds are recorded and described for the first time. The aim of the present study is to contribute to the resolution of the ITD/IPD dichotomy in sound lateralization by comparing these ERPs with the ERP to BBs of a pure tone.

To elicit and record cortical responses to BBs of complex sounds, the original description of BBs for sine tones must be generalized to cover also the sounds with several spectral components. The BB is described as a phenomenon experienced when two tones with slightly different frequencies are dichotically delivered to the ears. It is equally correct to use the term "slightly different periods" instead of "slightly different frequencies" in this definition. This alternative definition is used in the present study to design the BBs of the complex sounds employed (i.e., click train and recurrent noise). The periods of the sounds to the two ears are slightly shifted in opposite directions to maintain an interaural period difference, meaning a slight interaural difference of fundamental frequency. Actually, this is the way that the phenomenon of BBs is experienced with complex sounds in real life. For instance, when the same note simultaneously played with two strings of a musical instrument is slightly out of tune, acoustical beats are heard. Frequency differences between the fundamentals and between all the respective harmonics of the two tones are in proportion with their individual frequencies. This is exactly the way it is followed in the present study to create beating of complex sounds, with the only difference that the two sounds that are slightly out of tune are delivered dichotically to the ears to produce the BBs.

Materials and methods

Subjects

Sixteen volunteers with normal hearing (ten males) between 22 and 26 years of age (median: 24) participated in experiments and they were paid for their time. These subjects were 16 of the 19 participants who took part in a recent study of our group (Ungan et al. 2019). Initially, five of them were given an audiological test and their hearing thresholds were measured to be between 12 dB SPL and 28 dB SPL at 250 Hz. Their mean threshold was 22 dB SPL at 250 Hz and this level was referred as zero dB nHL to express the sound intensities of stimuli in dB nHL. All of the subjects participated in experiments were tested to have thresholds not higher than 6 dB nHL for 250 Hz and be able to perceive readily the binaural beats (BBs). Written informed consent of each subject to participate in the experiments was obtained and they were paid for their participation time. Procedures of the study conformed to World Medical Association Declaration of Helsinki and were approved by the Ethics Committee of Koc University.

Stimulus design

Three types of single-cycle BB stimuli, respectively, with a pure tone, a click train, and a Gaussian noise, were designed using Praat software (Boersma and Weenink 2013) with 40-kHz sample rate and they were saved as "wav" files.

Single-cycle binaural beats (BBs) of a 250-Hz pure tone

We used the single-cycle pure-tone-BB stimuli that were designed in our laboratory to record directly the ERPs to BB responses with no confound from monaural frequency modulation responses (Ungan et al. 2019). Base frequency of the sine tone was 250 Hz. Its 4-ms period covered with a good temporal resolution in the physiologically relevant ITD range, which is known to be around 690 µs (Moore and Moore 1971). In about one-third of the BB cycle, the ITD remained within the physiological range which extend form − 0.7 ms to + 0.7 ms corresponding to the leftmost and rightmost lateral positions. The tone frequency was increased by 1.6 Hz in one ear while decreasing it by 1.5 Hz in the other to create single-cycle BBs with an interaural frequency difference of ca. 3,1 Hz. These monaural frequency changes were below the perceptual threshold reported by Grose and Mamo (2012) as ca. 2.8 Hz. Yet, when presented dichotically, the same shifts readily created the percept of BBs. This interaural frequency difference corresponded to a BB cycle of 322 ms, which included in fact 80 periods of the slower tone to the right ear (channel 2) and 81 periods of the faster tone to the left ear (channel 1).

Single-cycle binaural beats (BBs) of dichotic click trains and noise

For click train stimuli, electrical pulses of 0.5-ms duration were used. These clicks were repeated with a base period of 4 ms (i.e., at a base rate of 250 clicks/s) in between BB stimuli. Click train-BBs were produced by shifting for 322 ms their periods to 4.025 ms and 3.975 ms in the left and right ears (channel 2 and channel 1), respectively.

For noise stimuli, a 4-ms-long Gaussian noise with $\mu = 0$ and $\sigma = 2$ was created using "randomGauss (μ, σ)" function of Praat software mentioned above, and this noise epoch was repeated with a base period of 4 ms. Noise-BBs were produced by omitting the last sample point of each of the eighty-one 4-ms epochs going to the left channel and repeating once the last point of each of the eighty 4-ms epochs going to the right channel, thus shifting for 322 ms the base period to 3.975 ms and 4.025 ms in the left and right ears (channel 1 and channel 2), respectively.

Stimulation

The single-cycle BBs were repeated every 2002 ms and, in between the single BB cycles, the base sounds (i.e., 250-Hz pure tone, 250/s click train, and recurrent 4-ms noise) were presented to both ears. Therefore, in recording event-related potentials (ERPs) to BBs, the stimulus onset asynchrony (SOA) was 2002 ms, including a 322-ms BB cycle and a 1680-ms period of base sound without any beat. At the start of a BB cycle, the fundamental components of the sound signals in the two ears were in counterphase, corresponding to an initial ITD of 2 ms. Therefore, during a BB cycle of 322 ms, the interaural time difference (ITD) for the dichotic tones started from 2 ms (i.e., beyond physiological sound lateralization range), then decreased to 1-ms lead for one ear, then pass from zero, then increased again to 1-ms lead for the opposite ear, and finally retuned to 2 ms. Parameters and timing of the single-cycle BB stimuli used in recording the ERPs to three BB types of dichotic sounds are given in Fig. 1. It should be noted that the sound channels in which the period of the base sound was shortened/prolonged to produce the BBs were identical for the sine and noise stimuli, but they were crossed for the click train stimuli.

Single-cycle BB stimuli designed and stored as wav files were played using a USB sound card (MUSE Mini DAC) driven under Matlab, and a pair of audiological insert earphones (Etymotic, EARTONE-3A) was used to deliver the sound stimuli to the ears. Intensity of the tone stimulus at 250-Hz base frequency was set to 70 dB nHL (93 dB SPL). SPL of the tone was determined by converting its rms voltage driving the earphone to dB SPL with the help of the technical information provided in the data sheet of the earphone which states that, for a 50 Ohm-earphone, a driving voltage of 200 mV (rms) produces 102.5 dB SPL in a HA-2 coupler. Peak amplitude of the driving sine wave was thus adjusted to 95 mV for a tone with 93 dB SPL. The amplitude of the click train stimuli was adjusted to 10 dB lower than that of the pure tone, because this wide-spectrum sound was found by the participants to be more disturbing than the pure-tone stimulus when they were presented with the same amplitude as the tone. This was probably because the large number of low-frequency harmonic components in this sound caused an increase in total spectral power. The peak-to-peak amplitude of the noise stimuli, however, was kept equal to that of the pure tone, because its harmonic components in the low-frequency range were relatively less strong, thus less disturbing compared to the click train. Sound intensities around 90 dB SPL are indeed considerably high levels at which some cross-talk between ears would be expected. However, this should not have been an issue in our experiments owing to the high interaural attenuation of the earphone given in its data sheet as 90 dB, 98 dB, 92 dB, and 83 dB at 250 Hz, 500 Hz, 750 Hz, and 1 kHz, respectively. All of the subjects readily perceived the BBs when the above explained pure tone, click train, and noise stimuli were presented dichotically, and they were not able to detect any auditory change when the sound to either one of the ears was blocked. Taking into account the observations of Kollmeier et al. (2008) who reported that the perceptual

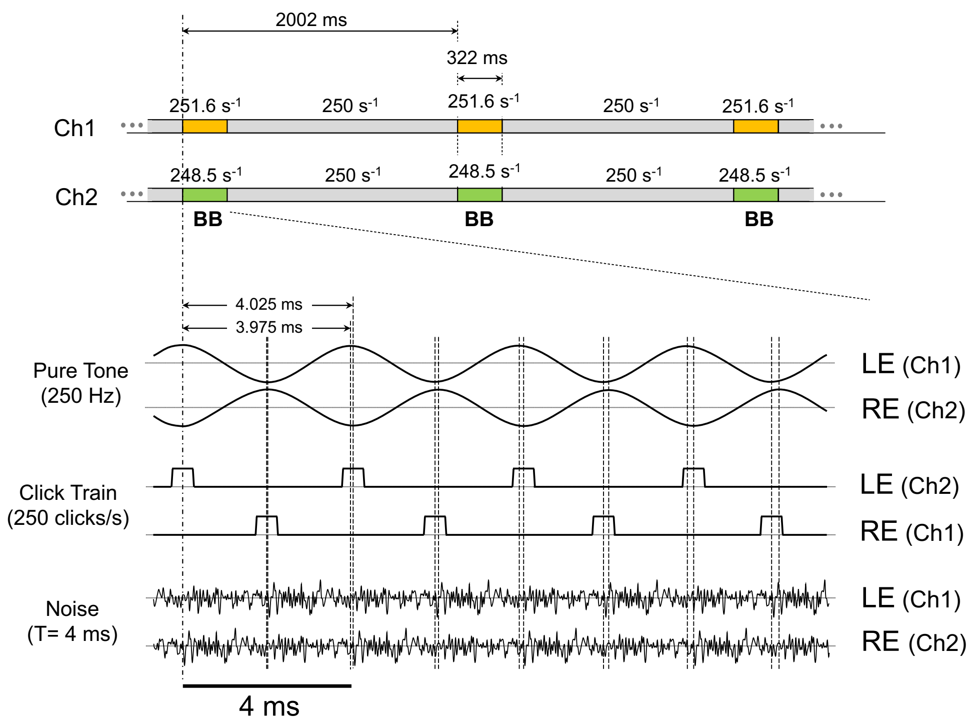

Fig. 1 Upper panel: parameters and timing of the two-channel stereo-signals to produce the single-cycle BB stimuli used in recording the ERPs to BBs of a dichotic tone with a frequency of 250 Hz, to a dichotic click train with a rate of 250 clicks/s, and to a dichotic noise recurrent at a rate of 250 repetitions/s. Lower panel: On an expanded time scale around the onset of a BB, the waveforms of the sounds that were dichotically presented to the ears to obtain the BBs. Please note that the period of the base sound was shortened in the left ear which is connected to channel 1 and prolonged in the right ear which is connected to channel 2 to produce the BBs of tone and noise, whereas these channels were swapped for click train stimuli. The earlier member of the vertical dashed line pairs indicates the middle or end points of the periods of the sound with higher frequency, whereas the later member indicates those points of the sound with lower frequency

threshold for perceiving a change (just noticeable difference; JND) in the frequency of a tone depends on the tone's frequency content and that the JND, which was about 3 Hz for sine waves below 500 Hz, may drop to 1 Hz for complex tones, we verified that the frequency shifts described above for the click train- and noise-BB stimuli were not perceived by any of the participants when either the left channel or the right channel of these stimuli was presented monaurally.

Waveforms of the electrical signals producing the sounds that were dichotically presented to the ears to obtain the BBs of a tone, a click train, and a noise that were used in the ERP experiments are given in Fig. 2, together with the amplitude frequency spectra of the sounds. The spectra are given only for frequencies up to 1.5 kHz, because BB processing is not expected beyond this frequency (Licklider et al. 1950). These sound spectra are obtained by modifying the FFT-computed spectral amplitude values of the electrical signal with the earphone frequency response provided by the manufacturer in the data sheet of the earphones for a Zwislocki Coupler. Actually, the sound spectra are similar to the spectra of the electrical signals producing them because of the fact that the earphone frequency response is essentially flat between 100 Hz and 1 kHz, with a 3 dB cut-off beyond 5 kHz. It is to be noted here that the earphone frequency responses in the data sheet are reported to be measured with a Zwislocki Coupler (Zwislocki 1971) and sound pressure in the Zwislocki coupler was shown to be essentially identical to pressure in real ears below 800 Hz (Sachs and Burkhard 1972). The same study reports also that, between 800 and 7500 Hz, the mean pressure in the real ears and in the Zwislocki coupler differs by not more than 2 dB.

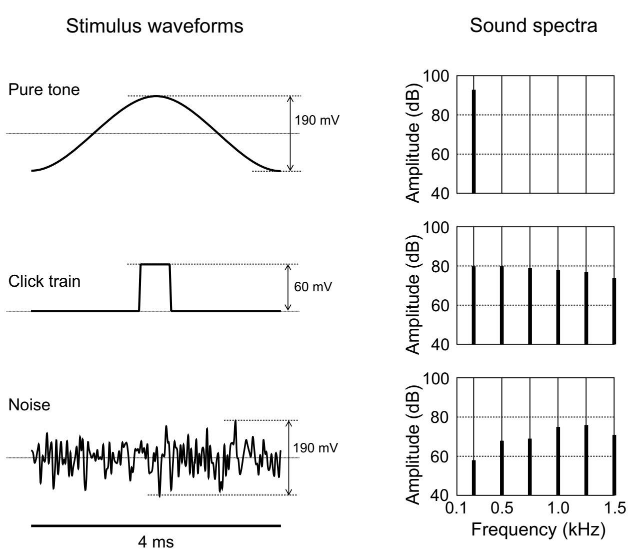

Fig. 2 Left panel: single-period waveforms of the three base signals used for driving the earphones to generate the BBs of a 250-Hz pure-tone, a 250/s click train, and a recurrent 4-ms noise. Right panel: The amplitude spectra of the sounds produced at the earphone's output by these electrical signals. These spectra are obtained by modifying the FFT-computed spectral amplitude values of the electrical signal with the earphone frequency response provided by the manufacturer of the earphones. The amplitude spectra are plotted up to 1.5 kHz considering the fact that the spectral components with frequencies beyond this frequency may not contribute to BB perception (Licklider et al. 1950)

Recording and analysis of ERPs

EEG recording with 32 channels was made by means of an electrode cap (EasyCap GmbH, Etterschlag, Germany) connected to the input of a 40-channel computerized recorder (Mega Electronics-NeurOne, Kuopio, Finland). Bandwidth of the amplifier and the sampling rate were set to 0.16–2500 Hz and 10 kHz, respectively. A high cut-off frequency and a high sampling rate were used, because the stimulation signals in the two sound channels were submitted to the amplifier's two auxiliary inputs and recorded simultaneously with EEG to align accurately the onset times of the BBs for temporal alignment of the ERP epochs to be averaged.

EEG was recorded simultaneously from 32 EEG channels, while the two electrodes slightly behind and slightly in front of the Fz electrode (REF and GND electrodes, respectively) served for referencing and grounding. In addition to the 31 scalp electrodes located according to 10/20 system, an additional electrode was attached to the nose tip for off-line re-referencing. Two of the extra channels of the system were used to record the stimulus signals from the left and right sound channels as mentioned above. Therefore, an unusually wide filter bandwidth of 0.16–2500 Hz and a very high sampling rate of 10 kHz were used.

During recordings, subjects sat on a reclining chair in an electrically and acoustically shielded chamber and they were instructed to minimize eye blinks and facial movements. Each subject participated in an ERP recording session with tone-BBs, during which 512 single sweeps were collected in four blocks, as well as in another session with click train-BBs and noise-BBs, during which 512 single sweeps were collected for each BB type. These sessions were conducted over 2 different days and the blocks were separated by 1–2 min for rest. Order of the sessions with tone and complex sounds was randomized. Subjects were instructed not to pay attention to stimuli and read a printed or electronic material of their choice during recording.

The raw EEG data were 0.5–30 Hz band-pass filtered and re-referenced to nose off-line. BB event triggers for temporal alignment of the ERP epochs to be averaged were added off-line. This was achieved by finding and counting the successive maximal peaks of the recorded stimulus tone, starting from its onset. The EEG was then epoched according to these BB event trigger times; and, using ± 100 µV artifact rejection, 600-ms single ERP epochs (including a 100 ms pre-stimulus part) were averaged for each BB type. Rejection rate was between 3 and 15% (mean 7%). Because the peak latencies of the ERP components varied over electrodes and the electrode at which a particular deflection reached a maximal amplitude displayed a notable intra subject variability (Scherg et al. 1989; Näätänen and Picton 1987), the global field power (GFP) curves (Lehmann and Skrandies 1980), rather than the studied ERPs themselves, were evaluated for reference-free assessment of the peak latency and dipolar strength of a certain response component. Nonetheless, the ERP waveforms were also inspected and provided. Peak latency of the N1 deflection in ERPs, which is the most characteristic auditory ERP component, was measured as the post-stimulus time of the GFP maximum corresponding to the maximum negative deflection at electrode Fz within the time range of 90–180 ms. The magnitude of the N1 response was quantified by the level of this GFP maximum, which should reflect the strength of the dipolar source of the N1.

Two-way ANOVAs were conducted to test if gender and the type of BB had a significant effect on the latency and magnitude of N1 in the ERPs recorded. Upon seeing that gender did not have a significant effect on these two dependent variables (see "Results" section), the data obtained from male and female subjects were pooled together in further statistical analyses. Post hoc *t* tests were conducted where necessary.

The scalp potential and surface current density distributions of the N1 deflection in the grand average ERPs of the subjects to single-cycle BBs were plotted in the form of 2D maps which were color-coded according to the mean of ± 10 ms around the N1 peak latency. In these surface current density maps, darkest red positive and darkest blue negative regions corresponded, respectively, to the current sources and current sinks (Tenke and Kayser 2012) associated with the N1 response. All steps of the above-mentioned off-line EEG and ERP processing were carried out using MATLAB and ERPLAB toolbox (Lopez-Calderon and Luck 2014).

Results

Grand average ERPs to single-cycle BBs of the 250-Hz tone, the 250-Hz click train, and the recurrent 4-ms noise, all recorded from 31 scalp electrodes and re-re-referenced to nose, are topographically presented in Fig. 3. In line with the typical characteristics of long-latency auditory-evoked potentials (Näätänen and Picton 1987), the ERPs to all three types of BB stimuli are dominated with an N1-P2 complex that is recorded with maximal amplitudes from fronto-central electrodes and display a polarity reversal at mastoids and occipital sites, especially in tone-BB responses. The P1, which is another typical auditory response component to be observed in the earlier latency range of 40–70 ms, could be recorded only in a few of the subjects and, therefore, it is rather obscured in the grand average ERPs. It is for this reason that the N1 response which is the most typical component of auditory ERPs is given special emphasis in the present study, and mainly the amplitude and latency of this deflection are evaluated.

The two-way ANOVAs conducted over the global field power (GFP) peak amplitudes and latencies of the N1 responses in ten male and six female participants showed that gender does not have a significant effect on these two parameters of N1 [Amplitudes: $F(1,42)=0.184$, $p=.67$; Latencies: $F(1,42)=0.934$, $p=.339$], and the interaction between gender and BB type is also not significant [amplitudes: $F(2,42)=0.85$, $p=.435$; latencies: $F(2,42)=0.008$, $p=.992$]. Therefore, data obtained from male and female subjects were pooled together in the following statistical analyses.

ERPs to tone-BB stimuli

ERPs to BBs of a 250-Hz tone recorded from 16 subjects and their grand average are presented in Fig. 4 (upper left). These ERPs to sine tone-BBs were part of the results of our recent study (Ungan et al. 2019), where they were compared with the ERPs to a temporary increase in tone amplitude to search for possible differences between the evoked responses of the auditory cortex involved in processing of the envelope and fine structure of a tone. In all the individual ERPs, which were recorded from the midline-frontal electrode (Fz) and re-referenced to nose, the N1 component is quite robust, whereas the P2 component is not so prominent in some of the subjects. Consequently, the P2 amplitude in the grand average ERP appears to be smaller than the N1 amplitude. The mean of the N1 amplitudes in individual ERPs to tone-BBs was computed to be 3.1 μV with a standard deviation of ± 1.1 μV. The N1 amplitude in the grand average ERP is slightly (10%) below this mean N1 amplitude, probably because the N1 latency had an inter-subject variability with a standard deviation of ± 13.6 ms around its mean value of 153 ms. GFP curves, which give the spatial standard deviation of scalp potential distribution as a function of post-stimulus time, are presented in Fig. 4 (lower left) for individual ERPs and their grand average. From these curves, the mean values of the GFP maxima around N1 and P2 latencies were calculated to be 1.5 ± 0.5 μV and 0.9 ± 0.4 μV, respectively. Prominence of N1 besides P2 is apparent also in these GFP curves. Scalp distributions of potential and surface current density of the grand average ERP around N1 latency are given in Fig. 5 (leftmost plots) as color-coded scalp maps.

ERPs to click train-BB and noise-BB stimuli

The individual ERPs recorded from 16 subjects to BBs of a 250/s click train and a recurrent 4-ms noise are presented,

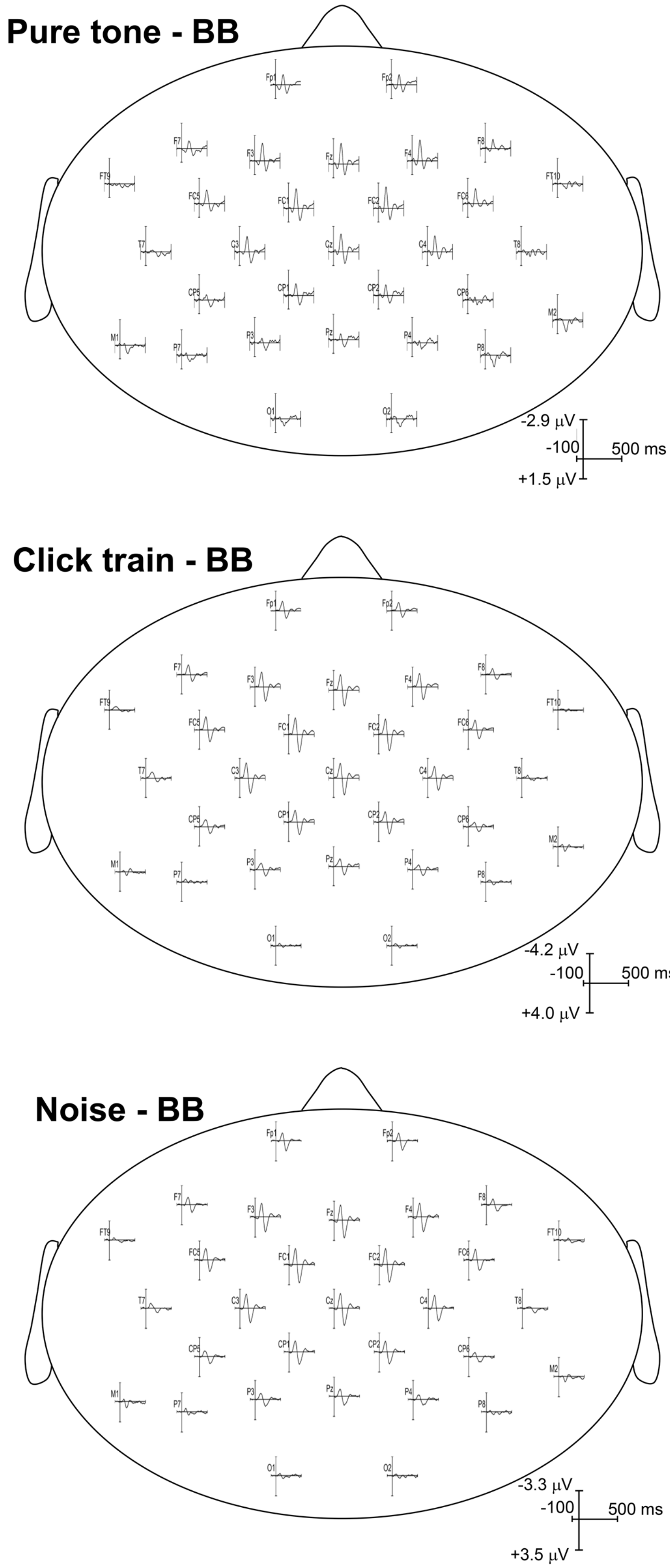

Fig. 3 Scalp topographical plots of the grand average ERPs recorded from 16 subjects in response to single-cycle BBs of a 250-Hz tone, a 250/s click train, and a recurrent 4-ms noise. ERPs are with nose reference

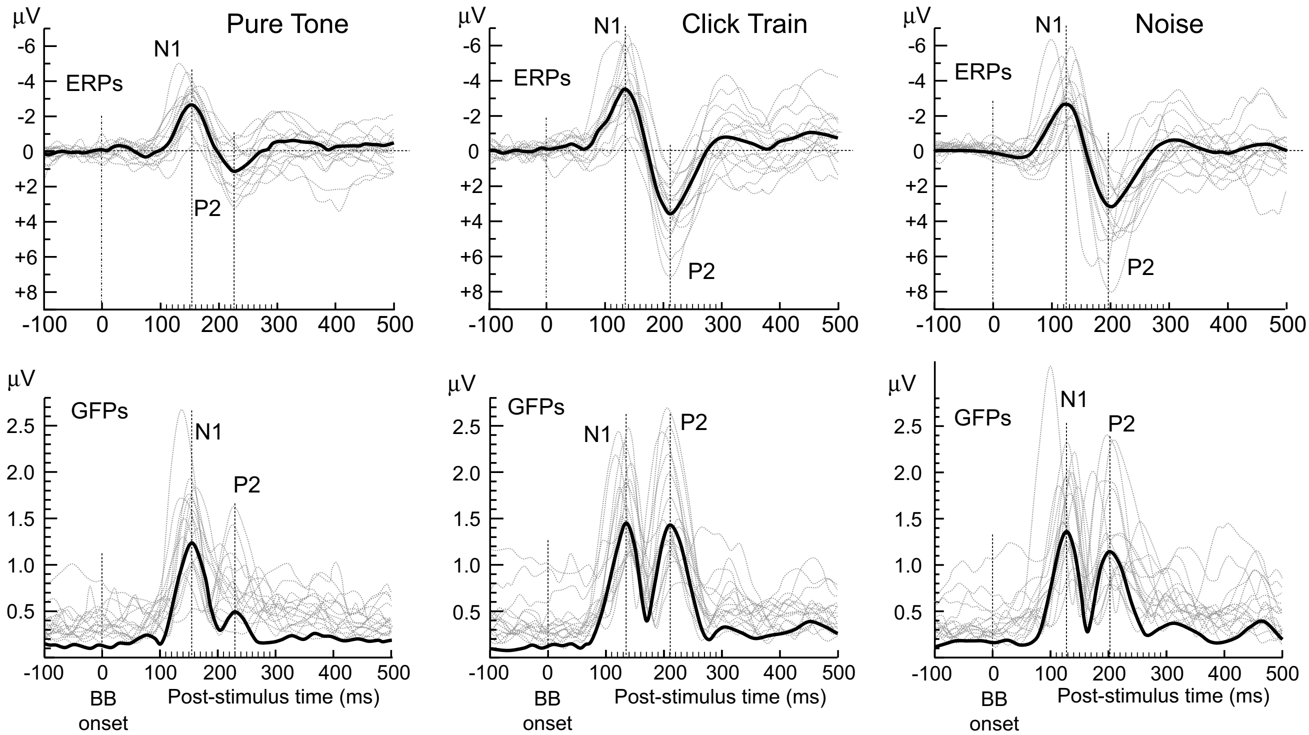

Fig. 4 Upper raw: superimposed individual ERPs of 16 subjects to BBs of a dichotic 250-Hz tone, of a 250/s click train, and of a recurrent 4-ms noise (thin lines) and their grand averages (heavy line). Recorded from the midline-frontal electrode (Fz) and re-referenced to nose. 100-ms pre-stimulus correction applied. Lower raw: GFP curves of these individual and grand average ERPs

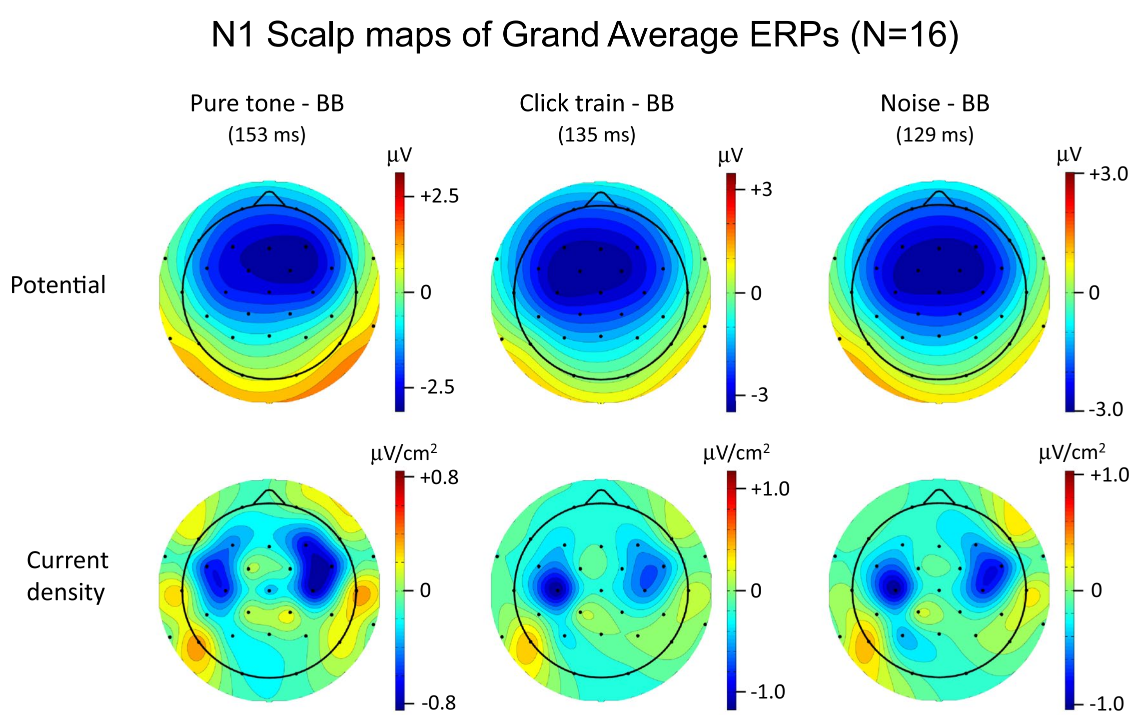

Fig. 5 Scalp surface potential and current density (Laplacian) maps of the N1 deflection in the grand average ERPs to single-cycle BBs of a 250-Hz tone, of a 250/s click train, and of a recurrent 4-ms Gaussian noise. Color-coded maps were plotted for the mean values over a ± 10 ms interval around the N1 latency given in parenthesis

respectively, in the upper middle and upper right of Fig. 4 together with their grand averages. GFP curves of these individual and grand average ERPs are also given in the lower middle and lower right of the same figure. P2 component is as robust as the N1 in most of the subjects. Consequently, these components in the grand average ERP appear to have comparable amplitudes. The N1 deflections evoked by the BBs of click train and noise had 3.9 µV and 3.4 µV mean amplitudes with standard deviations of ± 1.73 µV and 1.70 µV, respectively. The N1 amplitudes in the grand average ERPs are slightly below these mean N1 amplitudes, probably because the N1 latency had an inter-subject variability with standard deviations of ± 10.6 ms and ± 10.8 ms around their mean values of 135 ms and 139 ms for click train- and noise-BB stimuli, respectively. GFP curves of the ERPs to click- and noise-BBs, which give the spatial standard deviation of scalp potential distribution as a function of post-stimulus time, are presented in Fig. 4 (lower middle and right, respectively) for individual ERPs and their grand average. From the individual curves, the means of the GFP maxima of the ERPs to click- and noise-BBs around N1 latency were calculated to be 1.75 µV and 1.78 µV with standard deviations of ± 0.48 µV and ± 0.57 µV, respectively. The mean values of the GFP maxima around P2 latency were 1.62 ± 0.57 µV and 1.41 ± 0.52 µV, respectively. Compatibility of the N1 and P2 magnitudes is apparent also in these GFP curves. Scalp distributions of potential and surface current density of the grand average ERPs around their N1 latencies are given in Fig. 5 (middle and rightmost plots) as color-coded scalp maps.

Comparison of the ERPs

All of the 16 subjects were asked to compare the perceptual saliences of the BBs created by the tone with those of the click train and noise. They stated that the percept elicited by the beats of the latter two sounds was slightly but clearly stronger than that elicited by the beats of the tone. In line with this psychophysical observation, the amplitudes of the N1-P2 responses evoked in individual participants by the BBs of these two complex sounds were either comparable to or larger than the amplitudes of those evoked by tone-BBs, as can be seen by comparing the ERPs to click train- and noise-BBs with the ERPs to pure tone-BB in Fig. 4. Comparison of the N1 and P2 magnitudes in the respective GFP curves given in the lower raw of the same figure also indicates the superiority of the N1-P2 responses to complex sounds over the tone-BB-elicited N1-P2 responses. This superiority is especially evident for P2. Statistical test results supporting these observations are presented in Table 1 and Fig. 6.

Furthermore, the latency of N1 was shorter for click train- and noise-BBs than the tone-BB (on the average, 135 ms and 139 ms vs 153 ms, respectively). GFP and latency differences between the N1 s to BBs of complex sounds and those to 250-Hz pure tone proved to be statistically significant as shown by the results of paired t tests given in Table 1 and Fig. 6. However, neither the difference of 0.03 µV between the mean GFP amplitudes of the N1 responses to click train-BBs and noise-BBs nor the difference of 4 ms between their mean latencies was not found to be statistically significant [$t(15) = 0.3$, $p = .766$ and $t(15) = 1.61$, $p = .129$, respectively]. The amplitude and latency differences between the P2 responses to BBs of complex sounds, on the other hand, were statistically significant [$t(15) = 2.34$, $p = .034$ and $t(15) = 2.87$, $p = .008$, respectively].

Potential and surface current density maps of the N1 deflections seen in the grand average ERPs in Fig. 4 are given in Fig. 5. In the surface current density (Laplacian) maps (lower row), the positive regions colored in red correspond to current sources and the negative regions in blue correspond to current sinks (Tenke and Kayser 2012). There is no notable difference between the potential maps (upper row) of the N1 response to the click train- and noise-BBs.

Table 1 GFP amplitude and peak latency differences among the N1 and among the P2 responses to BBs of different sound stimuli and the results (Student's *t*, *p*, and Cohen's *d*) of the paired *t* tests run for assessing the significance of these differences

Comparison	N1 amplitude (GFP)			N1 latency		
	Difference	*t*	*p*	Difference	*t*	*p*
Tone vs click train	0.27 ($d = 0.72$)	2.88	.011	18 ($d = 2.22$)	8.9	.0000
Tone vs noise	0.3 ($d = 1.11$)	4.44	.0005	15 ($d = 1.39$)	5.57	.0000
Click train vs noise	0.03 ($d = 0.08$)	0.30	.77	4 ($d = 0.40$)	1.61	.13
Comparison	P2 amplitude (GFP)			P2 latency		
	Difference	*t*	*p*	Difference	*t*	*p*
Tone vs click train	0.75 ($d = 1.97$)	7.91	.000001	28 ($d = 0.97$)	3.81	.0017
Tone vs noise	0.54 ($d = 1.32$)	5.28	.00009	36 ($d = 1.13$)	4.59	.0004
Click train vs noise	0.21 ($d = 0.58$)	2.34	.034	9 ($d = 0.82$)	2.87	.008

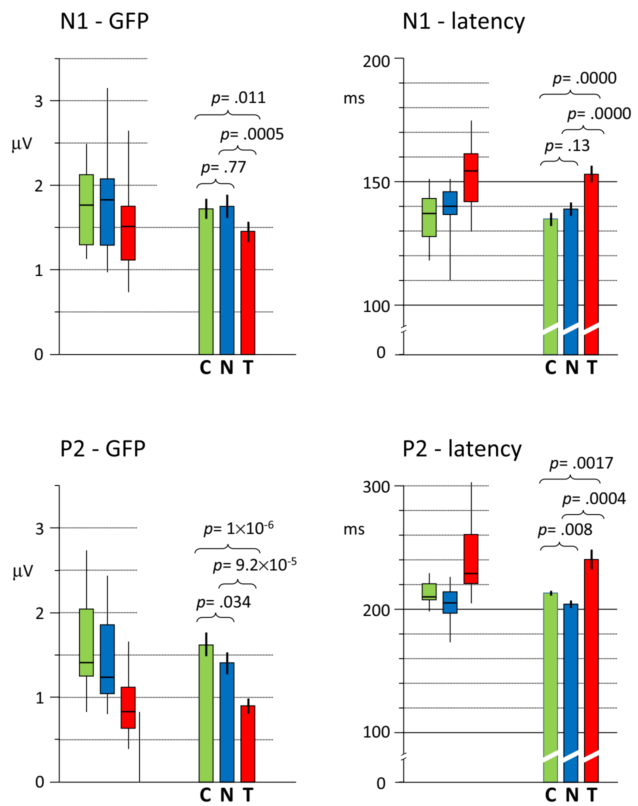

Fig. 6 Boxplot and bar graphics of the N1- and P2-GFP amplitudes and latencies in the ERPs to the three types of BB stimuli, summarizing the statistics of the data and significance of the amplitude and latency differences among the ERPs to the three BB types. Vertical lines over the mean value bars show the standard error of the mean. Letters C, N, and T under the bars denote the BB stimuli with click train, repetitive noise, and sine tone, respectively

Their current density maps also seem to be similar to each other. These two maps differ slightly from the N1 maps of the pure tone-BB response regarding the inter-hemispheric asymmetry they display. This slight spatial dissimilarity was not significant, however.

Discussion

Beginning and end of the BB cycle, where discontinuities in the slope of the change in ITD take place, may be the likely candidates of the events triggering the ERPs recorded. The N1 peak latency between 125 and 150 ms seems to fit with the onset of the BB cycle. There is no particular ERP deflection, however, at latencies which would match to its offset time. The absence of an offset response may be because the relatively short interval preceding the BB offset does not allow recovery of the responding cortical mechanism. Since we have no information about the exact phase of the BB cycle at which the ERP is triggered, we tend to assume that the whole BB cycle is processed as a unitary event, because the sound sequences falling within a window of integration of around 200–300 ms are believed to be coded holistically (Loveless et al. 1996).

The present study is based on the evaluation of the amplitude, latency, and scalp topography of the ERP-N1 response, which is the most typical component of auditory event-related potentials. Because this response is composed of several components which originate from distinct cortical sources having different dipolar orientations and recording distances, it is recorded with different amplitudes and peak latencies from different EEG electrodes (Scherg et al. 1989; Näätänen and Picton 1987). Because it is maximally recorded from the fronto-central scalp, it is common practice in N1 amplitude and latency evaluations that the recordings made from the Fz and/or Cz electrodes are considered (Näätänen and Picton 1987). However, the electrode from which this response is recorded with maximal amplitudes shows considerable inter-subject variability. Furthermore, amplitude and latency of N1 may also change depending on the location of reference electrode. In EEG recordings with relatively large number of electrodes, this variability can be avoided using a method called "global field power" (Lehmann and Skrandies 1980), which provides a reference-free measurement of amplitudes and latencies of ERP components by taking account the data collected from all electrodes. Global field power (GFP) is plotted against time, and the magnitude of a maximum in this curve is evaluated as the dipolar strength of the source associated with the ERP deflection having about the same latency with that GFP maximum. Because a component analysis was not intended in the present study, we used this approach and evaluated, in this sense, the maximum GFP value within the typical latency range of N1 as the amplitude of this response component. And, we also evaluated the post-stimulus time of that GFP peak as the latency of N1, compromising among various latency values that would have been read from different electrodes.

Of the three typical response components of auditory ERPs, namely P1, N1, and P2, the latter two were quite prominent in the ERPs to all three types of BB stimuli, as can be seen in Figs. 3, 4, and 7. P1, on the other hand, was either not distinct or totally absent. This may be explained by the fact that the amplitude of this deflection largely depends on the sharpness of the stimulus onset; i.e., its rise time (Kodera et al. 1979). None of the three BB stimuli used was in the form of an abrupt change; they were rather slow changes of interaural time or slow phase shifts with a relatively long cycle period of 322 ms. Therefore, our auditory stimuli did not have an adequately steep onset like a click sound or tone-pip, or an abrupt shift in sound frequency, which are known to be optimum for eliciting a prominent P1 response.

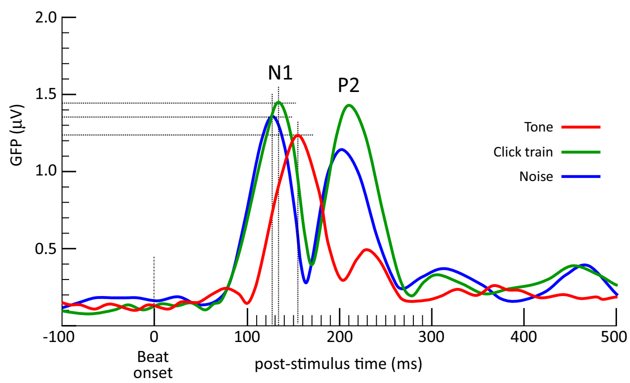

Fig. 7 Global field power plots of the grand average ERPs to binaural beats of a 250-Hz tone, of a 250/s click train, and of a recurrent 4-ms-long noise

While in ERPs to tone-BBs, the P2 component of the N1-P2 complex is much smaller in amplitude than the N1 component and it is even absent in some subjects, this positive wave appears as a prominent deflection in the ERPs to BBs of click trains and noise. There are studies showing that this positive wave may be related with cognitive processes such as selective attention (Hillyard et al. 1973), stimulus change and feature detection (Näätänen 1990; Luck and Hillyard 1995), short-term memory (Golob and Starr 2000), modulation of attention (Novak et al. 1992) and stimulus classification (Garcia-Larrea et al. 1992), and that its amplitude is enhanced by learning and attention (Crowley and Colrain 2004). However, because our subjects were not given any task concerning the above-mentioned cognitive functions during ERP recordings, the amplitude increase observed in P2 for click train- and noise-BBs cannot be connected to such effects. The results of a study (Shahin et al. 2005), which investigated whether auditory-evoked N1 and P2 responses are modulated by the spectral complexity of musical sounds in pianists and non-musicians, should be mentioned in connection with the enhanced P2 amplitude observed in the present study for BBs of click trains and noise. They reported that, not only in musicians but also in non-musicians, a piano tone with harmonics evoked a P2 with larger amplitude than the same tone without harmonics, although both stimuli were at the same sensation level (i.e., 60 dB above their respective sensory thresholds). This is a parallel observation to the one we made regarding the amplitude advantage of P2 for the two complex sounds we tried over the 250-Hz pure tone. However, in contrast to our results, a similar advantage was not observed for N1 in their study. Based on studies indicating parallels between the timing of onset responses of the single neurons in cat and of the evoked responses in human auditory cortex (Biermann and Heil 2000), Shahin et al. (2005) interpreted their finding to suggest that N1 is more sensitive to the parameters of temporal envelope of the stimuli, but P2 is more sensitive to their spectral content. This view may provide an explanation for the complexity-related N1-enhancement discrepancy between their study and the present one, considering the fact that the auditory event eliciting both types of BB responses (i.e., the response to BBs of a tone and to BBs of a complex sound) is not a change in the level or envelope of the sound, but a change in its ITD (i.e., its temporal fine structure). It is interesting, however, to note that the complexity-based enhancement of P2 in our study appears to be much greater than that of N1. It can clearly be seen in Fig. 6 that the complexity-related enhancement of N1 amplitude is 18% and 20% for click train- and noise-BBs, respectively, whereas the amplitude enhancement of P2 is 98% and 60% for the same stimuli (please also compare the N1 and P2 amplitudes of the grand average ERPs presented in Fig. 7). This finding indicates that, although the amplitudes of both N1 and P2 are enhanced by the spectral complexity of the stimulus when a change in its temporal fine structure is the auditory event eliciting the ERP, the enhancement in P2 is much larger than in N1.

There are a few works in the literature where ERPs to single-cycle BBs of pure tones were recorded and studied (Bohorquez and Ozdamar 2008; Özdamar et al. 2011; Ungan et al. 2019). However, we could find no study in which ERPs to BBs of complex sounds like click trains and noise are reported. In fact, two ERP studies of Ungan et al. (1989, 2001) show similarity to the present one, because the stimuli used in these earlier studies were also auditory lateralization shifts of click trains. In the present study, like in the earlier ones, the neural mechanisms stimulated are those detecting and processing the lateralization shifts produced by changing the interaural time difference (ITD) of dichotic click trains. The only difference is that the ITD was changed abruptly in the earlier studies, whereas in the present study, the ITD change is gradual and relatively slow according to the definition of the BB phenomenon. Therefore, the N1-P2 deflections in the ERPs reported by Ungan et al. (2001) and those presented in the present study should be the responses of basically the same cortical mechanisms. In principle, they have a certain similarity regarding their amplitudes, latencies, and scalp topographies. The amplitude and latency values reported by Ungan et al. (2001) are 4.7 µV and 124 ms for N1, and 5.7 µV and 206 ms for P2; whereas 3.9 µV and 135 ms, and 4.9 µV and 210 ms were measured in the present study for the same ERP components, respectively. The relatively smaller amplitudes and longer latencies measured in the present study may be explained by the difference between the rates (slopes) of ITD change in two cases. In the earlier study, a 1-ms change in the ITD of dichotic clicks was made in a period of 10 ms, whereas in the present study, 1-ms change in ITD, which corresponds to a quarter of the inter-click interval, takes 80 ms corresponding to a quarter of the BB cycle. In other words, the rate of ITD change was 1/10 in the earlier study and it is 1/80 in the present one. And, it is a rule of thumb in electrophysiology that the electrical response of a neural mechanism to a slowly changing stimulus (i.e., a stimulus with a long rise time or with a slope that is not very steep) is smaller in amplitude and longer in latency.

A right hemispheric dominance of N1-sink is notable in the Laplacian map of the tone-BB responses (Fig. 5, left column). It is probably because the frequency of the sound was increased in the left ear and decreased in the right ear for a duration of 322 ms to create single-cycle BBs, resulting in a right-to-left lateralization shift in the intracranial sound image. A dominance contralateral to the shift direction is also observed with the click train-BBs (Fig. 5, middle column) for which the rate of the clicks was decreased in the left ear and increased in the right ear (contrary to tone-BBs), resulting this time in a left-to-right shift of the sound image. In fact, a similar but less distinct contralateral dominance for the ITD-induced center-to-left shifts of a sound image is also seen in the grand average scalp map of the N1 response presented in the above-mentioned ERP study of Ungan et al. (2001). The scalp map of the N1 to noise-BBs does not seem to display any inter-hemispheric asymmetry, however (see Fig. 5, right column). It seems further experiments, in which ERPs to tone-BBs in both directions are to be recorded in the same session, would be needed for a decisive demonstration of such an inter-hemispheric asymmetry, and to see if it is due to a contralateral dominance related to the direction of lateralization shifts or whether it reflects a functional hemispheric dominance regardless of the shift direction.

To the best of our knowledge, there is only a single BB study that is conducted using a base sound with complex spectrum (Akeroyd 2010). And, this study on BBs of synthesized noise, which will be discussed below, is behavioral and does not include electrophysiological data. Therefore, the click train- and noise-evoked potentials reported in the present study are the first ERPs recorded with BBs of a sound with complex frequency spectrum. An immediate observation is that the BBs produced by dichotic presentation of a complex sound such as a click train or a noise also evoke cortical responses that are similar to those evoked by BBs of pure tones. Details of these new ERPs evoked by BBs of complex sounds will be given below in comparison with the ERPs to relatively familiar tone-BBs, and their differences will be discussed in connection with binaural sound lateralization mechanism.

In the behavioral study of Akeroyd (2010) in which BBs were produced by delivering to the left ear a synthesized noise with 400-Hz bandwidth centered around 500 Hz and to the right ear the same sound but its whole spectrum shifted upward or downward. Therefore, the large number of the harmonic Fourier components of the noise was all phase-shifted by the same amount. And, the subjects experienced a more powerful BB percept with this stimulus than a BB with a pure tone. This behavioral finding revealing an across-frequency integration of the phase-shift effects may suggest a mechanism based on interaural phase difference for BB perception. In the present study, however, the sounds in the two ears were shifted in time, and the phase shifts of their spectral Fourier components were not equal to each other, but proportional with the component frequency. And, all of the subjects perceived the BBs of click trains and noise at least as strongly as the BBs of the 250-Hz tone. This psychophysical observation is also reflected in the ERPs recorded with the BBs of these stimuli; the cortical responses to BBs of these complex sounds had at least as large an amplitude and as short a latency as the responses to BBs of a tone. N1 responses to click train- and noise-BBs had even larger global field powers and shorter latencies than those to tone-BBs (see the GFP curves given in Fig. 7 for the three types of BB stimuli for comparison).

These results show also electrophysiologically that the binaural beating effect which is described originally for sine tones can elicit cortical responses also when it is produced with complex sounds, and that this binaural phenomenon is perceived even more strongly for wide-band sounds having a number of spectral components. We tend to interpret this finding as indicating activation of larger number of ITD-sensitive neurons by the BBs of spectrally rich sounds than pure tones. However, before accepting this interpretation, one should consider the following two alternative explanations. The first explanation might be related to a well-known fact that stimuli with higher intensities elicit N1 responses with larger amplitudes and shorter latencies (Näätänen and Picton 1987). However, the event eliciting the N1 response in our study was not an onset or offset of a stimulus but a percept caused by a change in some feature of an ongoing sound. The parameter that should affect the amplitude of N1 is not, therefore, the intensity of the ongoing sound but the magnitude and rate (steepness) of the change; or, more accurately, the strength or salience of the percept caused by the change (Picton et al. 1974; Gutschalk et al. 2005). Therefore, the observed amplitude and latency advantage of the N1 responses to BBs of the complex sounds over BBs of the tone cannot be explained by a possible intensity advantage of the former stimuli. It should also be remembered here that the perceived intensities of the base sounds used for the three types of BBs were tried to be balanced by applying a 10 dB amplitude reduction to the peak-to-peak value of the click train waveform.

The second explanation for the amplitude differences observed between the click train-BB and pure tone-BB responses might be connected to the reversed sound channels in the two types of BB with respect to the direction of frequency change (see Fig. 1), which should lead to sound image movements in opposite directions. However, we observe the same complex sound advantage when we compare the amplitudes of the N1 s to noise-BBs and tone-BBs, which were not different in this respect. Furthermore, there was no significant amplitude difference between the N1 s evoked by click train-BB and noise-BB stimuli which had the same contrast regarding the direction of ITD change during the BB cycle, showing that a difference in the direction of possible sound image movements may not be an issue. We, therefore, consider the finding that the N1 responses in the ERPs to click train- and noise-BBs have larger amplitudes and shorter latencies than the N1 responses in the tone-evoked ERPs as an evidence for across-frequency integration of ITD-specific neuronal activation by the BBs of spectrally rich sounds.

Results of the present study provide supportive evidence for the view that a mechanism based on interaural time differences (ITDs), rather than interaural phase differences (IPDs), is responsible for BB perception. This view, which is in agreement with the conclusion reported by Zhang and Hartmann (2006), is based on the following argument: When dichotic sine tones are used for BBs, the time and phase differences between the sounds in the two ears (ITD and IPD) change together and linearly with time during a BB cycle of 322 ms. When BBs are produced by dichotic click trains; on the other hand, the ITDs of all sinusoidal harmonic components of the sounds in the two ears constantly change together during a BB cycle of 322 ms (Fig. 8, upper left, ITD vs time curve), whereas IPDs of the harmonic components differentiate

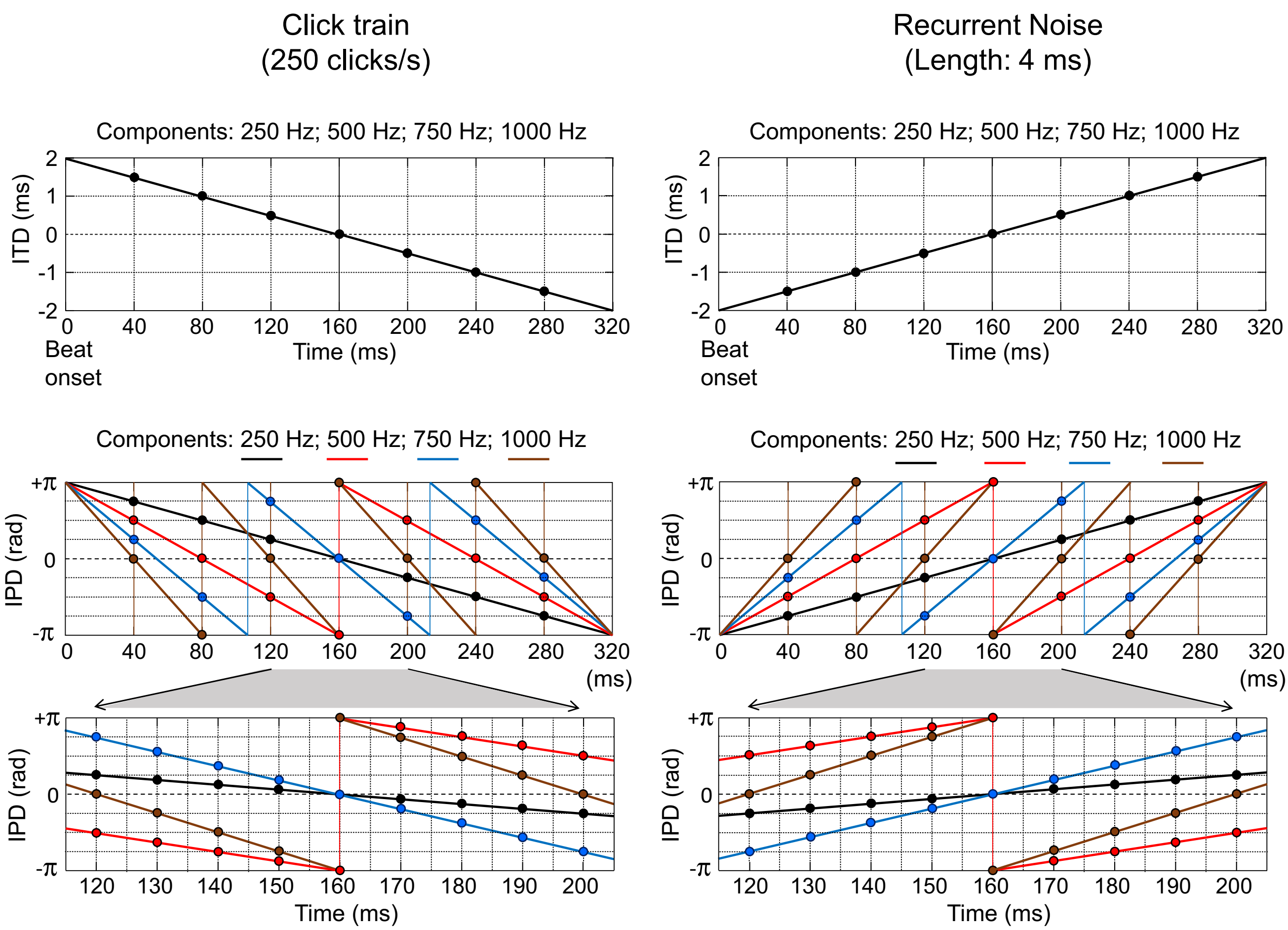

Fig. 8 For the BBs of a 250/s click train (left panel) and recurrent 4-ms-long noise (right panel), ITD vs time (upper curves), and IPD vs time plots (lower curves) of the fundamental and three harmonic components. It is seen that the interaural time differences (ITDs) of the fundamental component and all of its harmonics follow the same time course. Consequently, at any time point during the 322-ms beat cycle, there is a single ITD for the fundamental and all of its harmonic frequencies. However, the interaural phase difference (IPD) of each harmonic component follows a different time course. Consequently, at a certain time point following the BB onset, IPD takes different values for these harmonic frequencies, as demonstrated for different time points in the IPD vs time plots. The curves at the bottom illustrate this fact on an enlarged time scale to zoom in the 80 ms around the middle of the beat (i.e., between 120 ms and 200 ms), which is the time period of the beat cycle where ITD is within the physiological range between – 500 µs and + 500 µs

throughout the BB cycle, because their frequencies are different (Fig. 8, lower left, IPD vs time curves). In other words, the whole waveform, thus all of its harmonic components will shift in time together, but at any time point, each component will have a different IPD, as illustrated for 250 Hz and its first three harmonics at various time points after the BB onset). The case with the BBs of recurrent Gaussian noise can be described similarly as illustrated in right panel of the same figure.

Based on the arguments above, it is quite likely that the BB perception mechanism in the brain makes use of the ITD cue, which happens to be the same for all of the frequency components, rather than an IPD cue, which differs from one frequency component to another. As mentioned above, our observation that ERPs to BBs of click trains and noise were recorded with larger amplitudes than BBs to a tone suggests an advantage of complex sounds with many frequency components for the sound lateralization mechanism processing their BBs. This viewpoint is supported by a parallel observation that the latency of N1, which should be connected to perception time, was shorter for the BBs of wide-band sounds than for the sine tone-BB. And, such an advantage also speaks for an ITD-based mechanism, because only a binaural cue which is consistent across frequency can point to the same lateral position for all of the spectral components and may provide a perceptual advantage for the lateralization of wide-band stimuli.

It is true that neurons in the medial nucleus of superior olivary complex respond to their favorable IPD with spike rates significantly exceeding the sum of separate ipsilateral and contralateral stimulation (Batra et al. 1997), and this might suggest an IPD-based coding of binaural information in the brainstem. However, it was shown by Takahashi and Konishi (1986) that in some time-sensitive units of the inferior colliculus, simultaneous presentation of two tones with the same ITD produces greater responses than the sum of the responses to each of the tones presented in isolation. This finding, which is interpreted by Stern and Trahiotis (1992) as an interaction of second-level ITD coincidences across frequency, constitutes a physiological basis for recording stronger N1 responses to single-cycle BBs of ITD-shifted wide-band sounds with rich spectral composition than those with a single sine component (i.e., a pure tone). Furthermore, as the processing of auditory information continues in the tegmental and thalamic nuclei and also in the auditory cortex, the phase locking of neuronal activity to the tones decreases, but the interaural timing information is preserved (Kuwada et al. 2006). Since the ERP-N1 evoked by BBs is the response of the auditory cortex, as all the N1 deflections evoked by other auditory stimuli are (e.g., Scherg et al. 1989), our conclusion that ITD rather than IPD is utilized and processed in BB perception seems plausible also from this point of view.

The enhancement in the perceived intensity of the BB event observed in the study of Akeroyd (2010) where BBs were produced by shifting by the same amount the frequencies of all the spectral components of a noise is probably because several parallel phase-sensitive binaural units with different best frequencies in the superior olive were all activated synchronously with each other. The enhanced power of the percept may simply be due to recruitment of more units at the level of brainstem and not to any higher order across-frequency integration based on ITD consistency. In the present study, however, the spectral components of the complex sounds are each shifted in frequency by different but proportional amounts during BBs, and there may be no phase consistency and no synchronous activation of phase-sensitive units within the physiological range of interaural disparities that may be used for sound lateralization (shown on an expanded time scale in Fig. 8, bottom panel). The ITD information, on the other hand, is consistent over the spectral components of the complex sounds. The enhancement in BB salience as well as in ERP amplitudes that we observed should, therefore, be because of synchronous activation of the ITD-sensitive units with different best frequencies in some higher order centers such as tegmental and thalamic nuclei and also in the auditory cortex, where across-frequency integration of ITD-consistent neural activities may take place.

As a last point, let us discuss our results from the viewpoint of McFadden and Pasanen (1975), stating that there is an interaction in the nervous system between the envelopes extracted from the two ears. In other words, let us consider a model which evaluates the ITD of the envelope and not the ITDs of individual frequency components when the sound is a complex one. Actually, a BB model based on the envelope ITD is described and discussed by Bernstein and Trahiotis (1996) using an analogy between the envelope of a combination of high-frequency tones and a sine wave with a frequency equal to the fundamental frequency of the envelope. In a way, they presume that only the fundamental component of the sound is evaluated by the sound lateralization mechanism, ignoring the individual harmonic components of the sound waveform. However, such a model for BB perception would fail to explain our observation that larger N1-P2 amplitudes are recorded with BBs of multi-harmonic sounds compared to BBs of a sine tone. This is because the 250-Hz fundamental components of the click train and noise stimuli employed in the present study have amplitudes which are 13 dB and 35 dB lower than the amplitude of the pure-tone stimulus, respectively (please compare the amplitude spectra of the three sounds presented in Fig. 2). Recording larger responses to BBs of click trains or recurrent noise should indicate, therefore, contribution of the harmonic components in BB perception. Considering the fact that such wide-band stimuli possess ITDs that are consistent over a wide range of frequencies, this advantage of multi-harmonic sounds over a sine tone should be based on integration of consistent interaural timing information across frequencies; i.e., binaural processing of fine structure of the wide-band sound rather than merely its gross envelope repeating at its fundamental frequency. In the experiments of McFadden and Pasanen (1975), where two high-frequency tones were combined to form a low-frequency envelope, a periodicity pitch at that low frequency was perceived. However, there existed in fact no tone at that missing fundamental frequency in the BB-compatible low-frequency range, physically. An integration of consistent ITD information across multiple frequency components may not, therefore, occur in their BB experiments; and, the role of the time difference between monaural envelopes (envelope ITD) remains to be the only argument explaining their observations. It was probably for this reason that they evaluated the detectable beat with envelope periodicity as a faint one; much fainter than the conventional binaural beats of low-frequency tones. In the present study, on the other hand, the ERPs to BBs of complex sounds were recorded with larger amplitudes than those to BBs of a pure tone, suggesting that the neural mechanism that is responsible for the phenomenon of BBs evaluates the fine structure rather than or beside the envelope ITD of the dichotically presented sounds and, integrates consistent ITD information across their frequency components. In fact, the similarities between the frequency-dependent behaviors of the N1 responses to envelope ITD and to fine structure ITD suggest the existence of a common cortical code for the envelope and fine structure ITD, rather than a distinct ITD processing based on envelope information independent of the carrier (Ross 2018). In relation to the above-mentioned experiments of McFadden and Pasanen (1975), it seems that the physical existence of a low-frequency fundamental component corresponding to the envelope is important for a salient percept of BB. In case the fundamental component is accompanied, furthermore, by its neighboring harmonics, the BB percept becomes even more salient, like in the present study, indicating across-frequency integration of ITD-consistent neural activities.

Because it is more plausible to explain the lateralization of complex binaural stimuli by a model based on the consistency of the interaural cues presented over a wide range of frequencies (Stern and Trahiotis 1992; Saberi 1995) or the straightness of the trajectory of the cross-correlation peaks across frequency (Stern et al. 1988), the results of the present study indicate ITD rather than IPD as the cue that is used in human perception of auditory sidedness, besides interaural level difference (ILD).

Conclusion

In the present study, the ERPs to BBs of click trains and noise were recorded for the first time and with fairly high signal-to-noise ratio. The BBs of sounds with complex spectra evoked stronger cortical N1-P2 responses than the BBs of a pure tone. This finding has been evaluated as an evidence supporting the hypothesis that the cortical mechanism responsible for sound lateralization in humans is based, beside interaural level differences, on across-frequency integration of interaural differences of time (ITD) rather than those of phase (IPD). The results indicate also that BBs of complex sounds like click trains may be used as a more efficient auditory brain stimulation than BBs of pure tones for studies where beat stimulation is employed recently to benefit from its possible effects on EEG and some psychological processes (see the relevant citations in Introduction).

Acknowledgements This work was funded by Turkish Scientific and Technological Research Council, Ankara (Project: 114S492) and supported by Koc University, School of Medicine, Istanbul; and also by Science Academy, Istanbul, Turkey. Suha Yagcioglu, one of the co-authors of this study, passed away before submission of the manuscript. In recognition of his immense contribution to the study, we keep his name in the author list.

Compliance with ethical standards

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be regarded as a potential conflict of interest.

References

Akeroyd MA (2010) A binaural beat constructed from a noise. J Acoust Soc Am 128:3301–3304
Batra R, Kuwada S, Fitzpatrick DC (1997) Sensitivity to interaural temporal disparities of low- and high-frequency neurons in the superior olivary complex. I. Heterogeneity of responses. J. Neurophysiol 78:1222–1236
Beauchene C, Abaid N, Moran R, Diana RA, Leonessa A (2016) The effect of binaural beats on visuospatial working memory and cortical connectivity. PLoS One 11:e0166630
Becher AK, Höhne M, Axmacher N, Chaieb L, Elger CE, Fell J (2015) Intracranial electroencephalography power and phase synchronization changes during monaural and binaural beat stimulation. Eur J Neurosci 41:254–263
Bernstein LR, Trahiotis C (1996) Binaural beats at high frequencies: listeners' use of envelope-based interaural temporal and intensitive disparities. J Acoust Soc Am 99:1670–1679
Biermann S, Heil P (2000) Parallels between timing of onset responses of single neurons in cat and of evoked magnetic fields in human auditory cortex. J Neurophysiol 84:2426–2439
Boersma P, Weenink D (2013) Praat: doing phonetics by computer. [Computer program, Version 5.3.56]. Retrieved from http://www.praat.org

Bohorquez J, Ozdamar O (2008) Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. Clin Neurophysiol 119:2598–2607

Chaieb L, Wilpert EC, Hoppe C, Axmacher N, Fell J (2017) The impact of monaural beat stimulation on anxiety and cognition. Front Hum Neurosci 11:251

Chait M, Poeppel D, de Cheveigne´ A, Simon JZ (2005) Human auditory cortical processing of changes in interaural correlation. J Neurosci 14(25):8518–8527

Crowley KE, Colrain IM (2004) A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin Neurophysiol 115:732–744

Derner M, Chaieb L, Surges R, Staresina B, Fell J (2018) Modulation of item and source memory by auditory beat stimulation: a pilot study with intracranial EEG. Front Hum Neurosci 12:500

Edmonds BA, Krumbholz K (2014) Are interaural time and level differences represented by independent or integrated codes in the human auditory cortex? J Assoc Res Otolaryngol 15:103–114

Gao X, Cao H, Ming D, Qi H, Wang X, Wang X, Chen R, Zhou P (2014) Analysis of EEG activity in response to binaural beats with different frequencies. Int J Psychophysiol 94:399–406

Garcia-Argibay M, Santed MA, Reales JM (2019) Efficacy of binaural auditory beats in cognition, anxiety, and pain perception: a meta-analysis. Psychol Res 83:357–372

Garcia-Larrea L, Lukaszewicz AC, Mauguiere F (1992) Revisiting the oddball paradigm. Non-target vs. neutral stimuli and the evaluation of ERP attentional effects. Neuropsychologia 30:723–741

Golob EJ, Starr A (2000) Age-related qualitative differences in auditory cortical responses during short-term memory. Clin Neurophysiol 111:2234–2244

Grantham DW, Wightman FL (1978) Detectability of varying interaural temporal differences. J Acoust Soc Am 63:511–523

Grose JH, Mamo SK (2012) Electrophysiological measurement of binaural beats: effects of primary tone frequency and observer age. Ear Hear 33:187–194

Gutschalk A, Micheyl C, Melcher JR, Rupp A, Scherg M, Oxenham AJ (2005) Neuromagnetic correlates of streaming in human auditory cortex. J Neurosci 25:5382–5388

Halliday R, Callaway E (1978) Time shift evoked potentials (TSEPs): methods and basic results. Electroenceph Clin Neurophysiol 45:118–121

Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. Science 182:177–180

Hommel B, Sellaro R, Fischer R, Borg S, Colzato LS (2016) High-frequency binaural beats increase cognitive flexibility: evidence from dual-task crosstalk. Front Psychol 7:1287

Ioannou CI, Pereda E, Lindsen JP, Bhattacharya J (2015) Electrical brain responses to an auditory illusion and the impact of musical expertise. PLoS One 10:e0129486

Jones SJ, Pitman JR, Halliday AM (1991) Scalp potentials following sudden coherence and discoherence of binaural noise and change in the inter-aural time difference: a specific binaural evoked potential or a ‘mismatch’ response? Electroenceph Clin Neurophysiol 80:146–154

Knight RT, Scabini D, Woods DL, Clayworth CC (1988) The effects of lesions of superior temporal gyrus and inferior parietal lobe on temporal and vertex components of the human AEP. Electroenceph Clin Neurophysiol 70:499–508

Kodera K, Hink RF, Yamada O, Suzuki JI (1979) Effects of rise time on simultaneously recorded auditory-evoked potentials from the early, middle and late ranges. Audiology 18:395–402

Kollmeier B, Brand T, Meyer B (2008) Perception of speech and sound. In: Benesty J, Sondhi MM, Huang Y (eds) Springer handbook of speech processing. Springer, Berlin, p 65

Kuhn GF (1979) The pressure transformation from a diffuse sound field to the external ear and to the body and head surface. J Acoust Soc Am 65:991–1000

Kuwada S, Fitzpatrick DC, Batra R, Ostapoff EM (2006) Sensitivity to interaural time differences in the dorsal nucleus of the lateral lemniscus of the unanesthetized rabbit: Comparison with other structures. J Neurophysiol 95:1309–1322

Lehmann D, Skrandies W (1980) Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr Clin Neurophysiol 48:609–621

Licklider JCR, Webster JC, Hedlun JM (1950) On the frequency limits of binaural beats. J Acoust Soc Am 22:468–473

Lopez-Calderon J, Luck SJ (2014) ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front Hum Neurosci 8:213

Loveless N, Levänen S, Jousmäki V, Sams M, Hari R (1996) Temporal integration in auditory sensory memory: neuromagnetic evidence. Electroencephalogr Clin Neurophysiol 100:220–228

Luck SJ, Hillyard SA (1995) The role of attention in feature detection and conjunction discrimination: an electrophysiological analysis. Int J Neurosci 80:281–297

McAlpine D, Jiang D, Palmer AR (2001) A neural code for low-frequency sound localization in mammals. Nat Neurosci 4:396–401

McEvoy LK, Picton TW, Champagne SC, Kellet AJC, Kelly JB (1990) Human evoked potentials to shifts in the lateralization of a noise. Audiology 29:163–180

McFadden D, Pasanen EG (1975) Binaural beats at high frequencies. Science 190(4212):394–396

Moore JK, Moore RY (1971) A comparative study of the superior olivary complex in the primate brain. Folia Primatol 16:35–51

Näätänen R (1990) The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. Behav Brain Sci 13:201–233

Näätänen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24:375–425

Novak G, Ritter W, Vaughan HG Jr. (1992) Mismatch detection and the latency of temporal judgements. Psychophysiology 29:398–411

Özdamar Ö, Bohorquez J, Mihajloski T, Yavuz E, Lachowska M (2011) Auditory evoked responses to binaural beat illusion: Stimulus generation and the derivation of the binaural interaction component (BIC). Conf Proc IEEE Eng Med Biol Soc 2011:830–833

Papanicolaou A, Baumann S, Rogers R, Saydjari C, Amparo E, Eisenberg H (1990) Localization of auditory response sources using magnetoencephalography and magnetic resonance imaging. Arch Neurol 47:33–37

Perrott DR, Musicant AD (1977) Rotating tones and binaural beats. J Acoust Soc Am 61:1288–1292

Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials: I. Evaluation of components. Electroencephalogr Clin Neurophysiol 36:179–190

Pratt H, Starr A, Michalewski HJ, Dimitrijevic A, Bleich N, Mittelman N (2009) Cortical evoked potentials to an auditory illusion: binaural beats. Clin Neurophysiol 120:1514–1524

Pratt H, Starr A, Michalewski HJ, Dimitrijevic A, Bleich N, Mittelman N (2010) A comparison of auditory evoked potentials to acoustic beats and to binaural beats. Hear Res 262:34–44

Rayleigh L (1907) On our perception of sound direction. Phil Mag 13:214–232

Ross B (2018) Auditory cortex responses to interaural time differences in the envelope of low-frequency sound, recorded with MEG in young and older listeners. Hear Res 370:22–39

Saberi K (1995) Lateralization of comodulated complex waveforms. J Acoust Soc Am 98:3146–3156

Sachs RM, Burkhard MD (1972) Zwislocki Coupler. Evaluation with insert earphones. Report 20022. Industrial Research Products, Inc., Elk Grove, IL

Sams M, Hämäläinen M, Hari R, McEvoy L (1993) Human auditory cortical mechanisms of sound lateralization. I. Interaural time differences within sound. Hear Res 67:89–97

Scherg M, Vajsar J, Picton TW (1989) A source analysis of the late human auditory evoked potentials. J Cognit Neurosci 1:336–355

Shackletona TM, Meddis R, Hewitt MJ (1992) Across frequency integration in a model of lateralization. J Acoust Soc Am 91:2276–2279

Shahin A, Roberts L, Pantev C, Trainor LJ, Ross B (2005) Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport 16:1781–1785

Stern RM, Trahiotis C (1992) The role of consistency of interaural timing over frequency in binaural lateralization. In: Cazals Y, Horner K, Demany L (eds) Auditory physiology and perception. Pergamon Press, Oxford, pp 547–553

Stern RM, Zeiberg AS, Trahiotis C (1988) Lateralization of complex binaural stimuli: a weighted-image model. J Acoust Soc Am 84:156–165

Stewart GW (1917) Binaural beats. Phys Rev 9:502–508

Takahashi TA, Konishi M (1986) Selectivity for interaural time differences in the owl's midbrain. J Neurosci 6:3413–3422

Tenke CE, Kayser J (2012) Generator localization by current source density (CSD): implications of volume conduction and field closure at intracranial and scalp resolutions. Clin Neurophysiol 123:2328–2345

Ungan P, Yagcioglu S (2002) Origin of the binaural interaction component in wave P4 of the short-latency auditory evoked potentials in the cat: Evaluation of serial depth recordings from the brainstem. Hear Res 167:81–101

Ungan P, Şahinoğlu B, Utkuçal R (1989) Human laterality reversal auditory evoked potentials. Electroencephalogr Clin Neurophysiol 73:306–321

Ungan P, Yagcioglu S, Göksoy C (2001) Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. Clin Neurophysiol 112:485–498

Ungan P, Yagcioglu S, Ayik E (2019) Event-related potentials to single-cycle binaural beats and diotic amplitude modulation of a tone. Exp Brain Res 237:1931–1945

Vaughan HG Jr, Ritter W (1970) The sources of auditory evoked responses recorded from the human scalp. Electroencephalogr Clin Neurophysiol 28:360–367

von Wedel H (1982) Evoked cortical potentials caused by short-term changes in interaural coherence. Laryngol Rhinol Otol 61:159–164

Wernick JS, Starr A (1968) Binaural interaction in the superior olivary complex of the cat: An analysis of field potentials evoked by binaural-beat stimuli. J Neurophysiol 31:428–441

Yost WA (1981) Lateral position of sinusoids presented with interaural intensive and temporal differences. J Acoust Soc Am 70:397–409

Zhang PX, Hartmann WM (2006) Lateralization of sine tones-interaural time vs phase. J Acoust Soc Am 120:3471–3474

Zwislocki JJ (1971) An earlike coupler for earphone calibration. LSC-S-9, Laboratory for Sensory Communication, Syracuse Univ., Syracuse