Psychophysiological Evidence of an Autocorrelation Mechanism in the Human Auditory System

Yoshiharu Soeta

Additional information is available at the end of the chapter

Abstract

This article details a model for evaluations of sound quality in the human auditory system. The model includes an autocorrelation function (ACF) mechanism. Thus, we conducted physiological and psychological experiments to search for evidence of the ACF mechanism in the human auditory system. To evaluate physiological responses related to the peak amplitude of the ACF of an auditory signal, which represents the degree of temporal regularity of the sound, we used magnetoencephalography (MEG) to record auditory evoked fields (AEFs). To evaluate psychological responses related to the envelope of the ACF of an auditory signal, which is a measure of the repetitive features of an auditory signal, we examined perceptions of loudness and annoyance. The results of the MEG experiments showed that the amplitude of the N1m, which is found above the left and right temporal lobes around 100 ms after stimulus onset, was a function of the peak amplitude and its delay time or the degree of envelope decay of the ACF. The results of the psychological experiments indicated that loudness and annoyance increased for sounds with envelope decay of the ACF in a certain range. These results suggest that an autocorrelation mechanism exists in the human auditory system.

Keywords: auditory evoked field, pitch strength, loudness, annoyance

1. Introduction

Correlation is one of the most common and useful statistical concepts. It measures the strength and direction of a linear relationship between two variables. **Figure 1** shows some examples of correlations between pairs of variables, including white noise signals with different phases, pure tones with the same frequency and phase, pure tones with different frequencies, human

voice signals and time-delayed versions of the same signal, environmental noise signals and time-delayed versions of the same signal, and environmental noise signals obtained at the left and right ears. The correlation coefficient ranges between −1 and 1, and characterizes the strength of the relationships between the two variables.
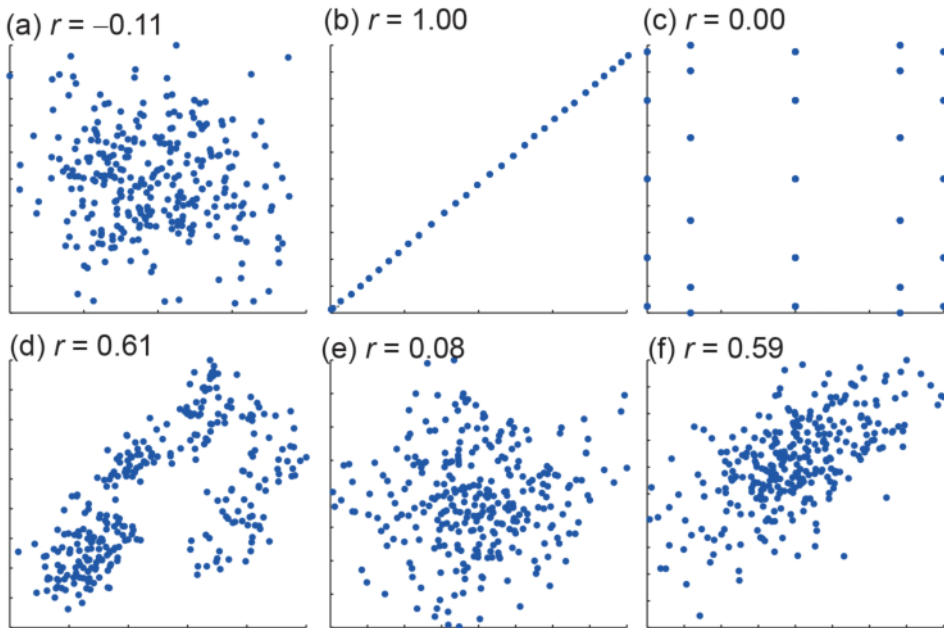

Figure 1. Relationship between two variables. (a) White noise signals with different phases, (b) pure tones with the same frequency, (c) pure tones with different frequencies, (d) human voice signals and time-delayed versions of the same signal, (e) environmental noise and time-delayed versions of the same signal, and (f) environmental noise signals obtained at the left and right ears.

When a signal is represented as a time series, it is characterized by periodicity or randomness as a function of time. **Figure 2** shows some examples of relationships between a signal and the time-delayed version of that signal. The signals included in the figure are white noise, pure tones, a human voice, and train noise. The way in which correlation coefficients change as a function of time can be evaluated using an autocorrelation function (ACF). An ACF is a set of correlation coefficients that characterize the relations between the points in a series and time-delayed version of the same set. In other words, the ACF is a time-domain function that measures how much a waveform resembles the delayed version of itself. While the values of an ACF can extend beyond −1 and 1, the normalized ACF (NACF) for a signal, $\phi(\tau)$, is defined by
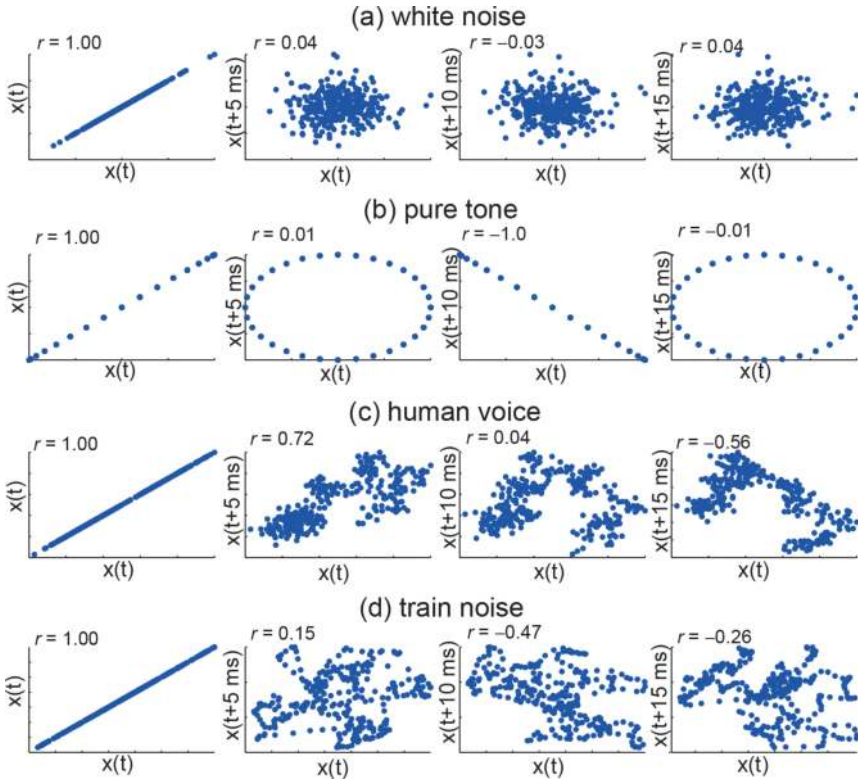
$$\phi(\tau) = \frac{\Phi(\tau)}{\Phi(0)} \tag{1}$$

Figure 2. Relationships between a signal and a time-delayed version of the same signal. (a) White noise, (b) pure tones, (c) a human voice, and (d) train noise.

where

$$\Phi(\tau) = \frac{1}{2T} \int_{-T}^{+T} p(t)\,p(t+\tau)dt \tag{2}$$

That is, the ACF is normalized by the maximum value of the ACF at the point with zero delay, $\Phi(0)$, thus restricting the values to fit the range between −1 and 1. **Figure 3** shows some examples of the NACF. As white noise is random, the ACF is close to zero. As pure tones are completely periodic, the ACF is also periodic and the maximum and minimum values are 1 and −1, respectively. The human voice and environmental noise have periodic components, so the ACF values for these stimuli are high at the dominant frequency.
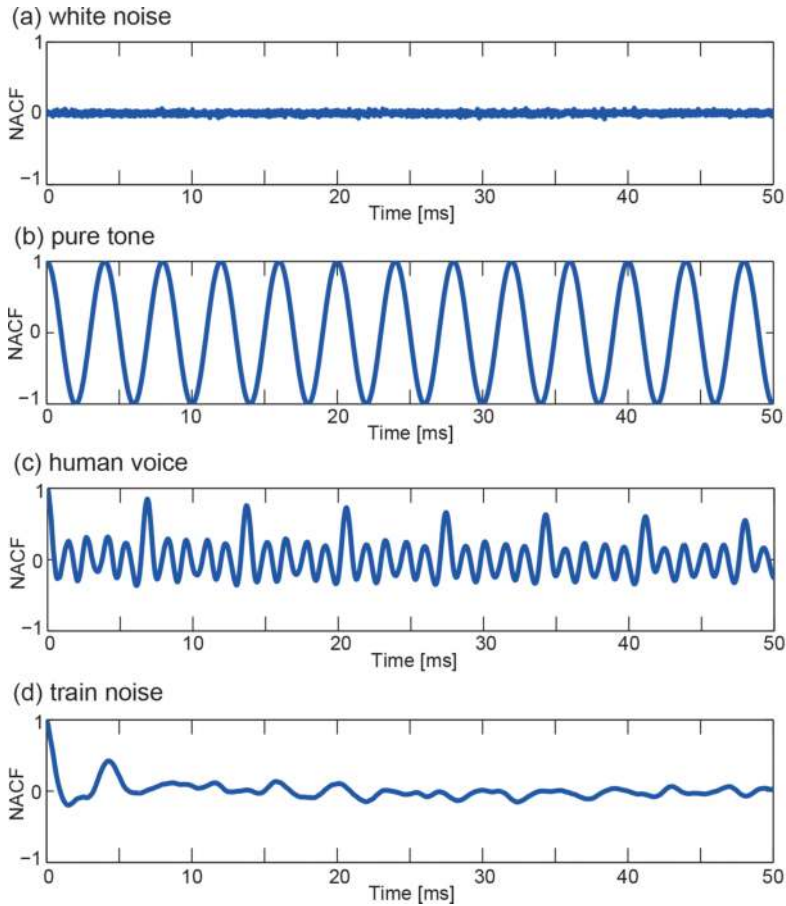
Figure 3. Examples of the NACF for (a) white noise, (b) pure tones, (c) the human voice, and (d) train noise.

Mathematically, the ACF contains the same information as the power spectrum of a given signal. For characterization of auditory signals, five factors are extracted from the ACF [1]. The first factor is the energy at the point with zero delay, given by $\Phi(0)$, which corresponds to the equivalent continuous sound pressure level (SPL). The second and third factors are the amplitude and delay time of the first maximum peak of the NACF, ϕ_1 and τ_1, which are related to the perceived pitch strength and pitch [2, 3]. The fourth factor is the effective duration of the envelope of the NACF, τ_e, which is defined by the 10th percentile delay. It represents a repetitive feature containing the auditory signal itself and is related to the preferred condition for the temporal factors of a sound field, such as reverberation time and the delay time of the first reflection [3, 4]. The fifth factor is the width of the amplitude of the NACF around the origin of the delay time, $W_{\phi(0)}$, which is defined as having a value of 0.5. It corresponds to the spectral centroid [1]. The definitions of the ACF factors are depicted in **Figure 4**.
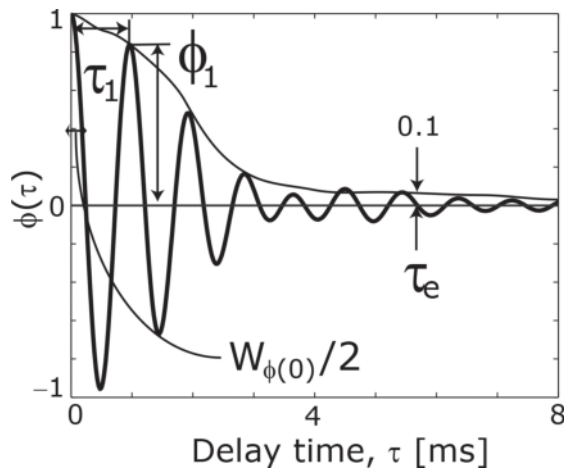
Figure 4. Definitions of the ACF factors, ϕ_1, τ_1, τ_e, and $W_{\phi(0)}$.

The ACF is one of the most famous models for describing the perception of pitch and pitch strength. Pitch is thought to be extracted by the ACF in the temporal model of pitch perception [e.g., 5–7] and pitch strength corresponds to ϕ_1, which represents the degree of temporal regularity of a sound [e.g., 1–3, 6]. It is possible to systematically manipulate the values of ϕ_1 using iterated rippled noise (IRN). IRN is produced by adding a delayed version of a noise signal to the original signal, and then repeating this delay and addition process [2]. Increasing the number of iterations increases the periodicity and ϕ_1 value.

Physiologically, IRN elicits signals in auditory nerve fibers [8, 9] and cochlear nucleus neurons [10–12], indicating that the pitch of IRN is represented in the firing patterns of action potentials locked to either the temporal fine structure or the envelope periodicity. That is, autocorrelation-like behavior in the fine structure of the neural firing patterns suggests that the pitch of IRN is based on an ACF mechanism. Indeed, the pooled interspike interval distributions of auditory nerve discharge patterns in response to complex sounds are similar to the ACF of the stimulus waveform, and ϕ_1 of the ACF corresponds to pitch strength [13, 14].

Therefore, to find the physiological counterparts of an ACF mechanism in the human auditory cortex, we used magnetoencephalography (MEG) to investigate the auditory evoked magnetic field (AEF) elicited by IRN and bandpass filtered noise (BPN). The ϕ_1 value can be manipulated systematically by changing the bandwidth of the BPN. A narrower bandwidth produces a higher ϕ_1. In MEG, the measured signals are generated by synchronized neuronal activity in the human brain. The time resolution is in the range of milliseconds. Thus, this technique can be used to examine rapid changes in cortical activity that reflects ongoing signal processing in the brain; electrical events in single neurons typically last from one to several tens of milliseconds [15]. With respect to the psychological aspect of sound perception, we evaluated the effects of the other ACF factor, i.e., τ_e, on loudness and annoyance because it can explain changes in loudness even when SPL conditions are unchanged.

2. AEFs in relation to the peak amplitude of the ACF, ϕ_1

2.1. AEFs in relation to IRN

MEG has been used to investigate how features of sound stimuli related to pitch are represented in the human auditory cortex. For instance, tonotopic organization of the human auditory cortex has been investigated as a spatial representation of pure tone in the auditory system according to frequency [16–18]. The frequency of pure tones has been found to influence the source location of AEF response components, such as the N1m, in the human auditory cortex. The periodicity of pitch-related cortical responses has been investigated as part of the temporal structure of sound [19, 20]. However, it is currently unclear whether periodic pitch is reflected in the location of the source of the AEF response in the human auditory cortex.
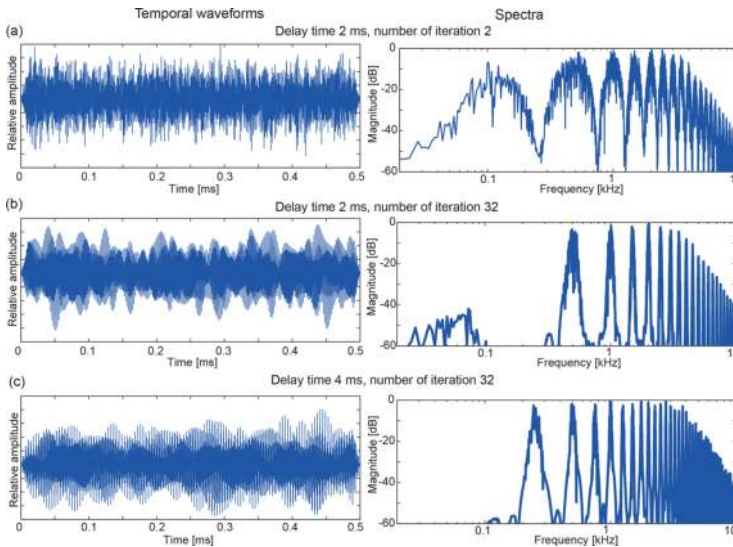

Figure 5. Temporal waveforms (left panels) and power spectra (right panels) of the IRN with different delay times (*d*) and number of iterations (*n*). (a) *d* = 2 ms, *n* = 2; (b) *d* = 2 ms, *n* = 32; (c) *d* = 4 ms, *n* = 32.

To evaluate responses related to the first maximum peak of the ACF, ϕ_1, which corresponds to pitch strength, in the auditory cortex, we recorded the AEFs elicited by IRNs with different iteration numbers. We anticipated that the N1m amplitude would increase with ϕ_1. The N1m is a typical component of the AEFs, which is generated in the auditory cortex approximately 100 ms after stimulus onset, offset, or a change in sound [21]. A large number of physical and psychological parameters have been reported to influence N1m responses, including intensity, frequency, interaural level or time difference, threshold, states of arousal, and selective attention. For example, the N1m is correlated with basic sensations such as loudness and pitch [1].

Ten normal-hearing listeners (22–36 years; all right-handed) took part in the experiment. We produced an IRN using a delay-and-add algorithm applied to BPN that was filtered using fourth-order Butterworth filters between 100 and 3500 Hz. The number of iterations of the delay-and-add process was set at 2, 4, 8, 16, and 32, and the delay was set to 2 and 4 ms, corresponding to pitch values of 500 and 250 Hz, respectively. The stimulus duration was 0.5 s, including rise and fall ramps of 10 ms. The sounds were digital-to-analog (D/A) converted with a 16-bit sound card and a sampling rate of 48 kHz. Sounds were presented at a SPL of 60 dB through insert earphones inserted into both the left and right ear canals. **Figure 5** shows the temporal waveforms and the power spectra of some of the IRN used in this experiment. **Figure 6** shows the ACF waveform of some of the IRN used in this experiment. The τ_1 value of IRN is the same value with the delay of the IRN. The ϕ_1 value increases as the number of iterations increases.
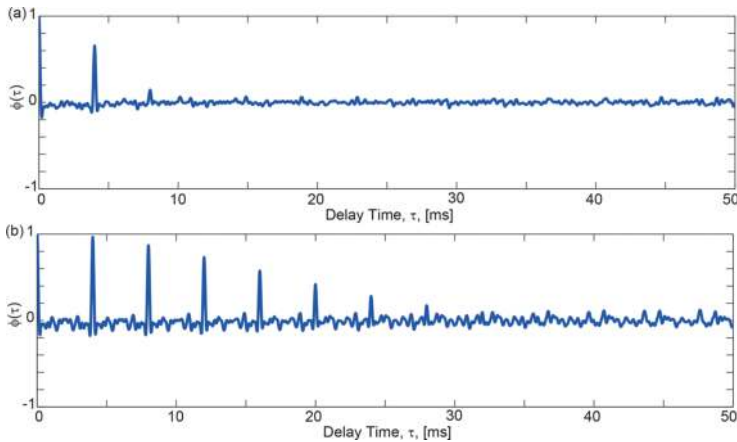

Figure 6. ACFs of the IRN with the delay time of 4 ms and number of the iterations: (a) 2 and (b) 32.

The AEFs were recorded using a 122 channel whole-head DC superconducting quantum interference device (DC-SQUID) magnetometer (Neuromag-122™; Neuromag Ltd., Helsinki, Finland) in a magnetically shielded room [15]. The IRNs were presented in a randomized order with a constant interstimulus interval of 1.5 s. To maintain listeners' attention level, listeners were instructed to watch a self-selected silent movie and ignore the stimuli during the experiment. The magnetic data were sampled at 0.4 kHz after being bandpass filtered between 0.03 and 100 Hz, then averaged approximately 100 times. The averaged responses were digitally filtered between 1.0 and 30.0 Hz. We analyzed a 0.7 s period starting 0.2 s prior to the stimulus onset, and an averaged 0.2 s prestimulus period served as the baseline.

We conducted source analysis for the measured field distribution based on the model of a single moving equivalent current dipole (ECD) [15]. Source estimates were based on a subset of 40–44 channels over each hemisphere. The dipole with the maximal goodness-of-fit over the analysis time window was chosen for further analysis. Only dipoles with a goodness-of-fit of

more than 80% were included in the further analyses. The source waveforms for all stimuli were calculated using the best-fitting dipole in each hemisphere. The peak amplitudes and latencies of the N1m reported in the following sections are based on the source waveforms.
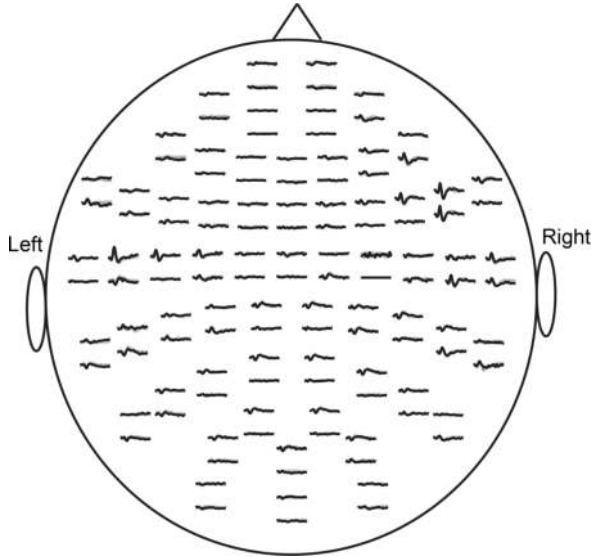

Figure 7. Typical waveforms of AEFs from 122 channels in a listener.
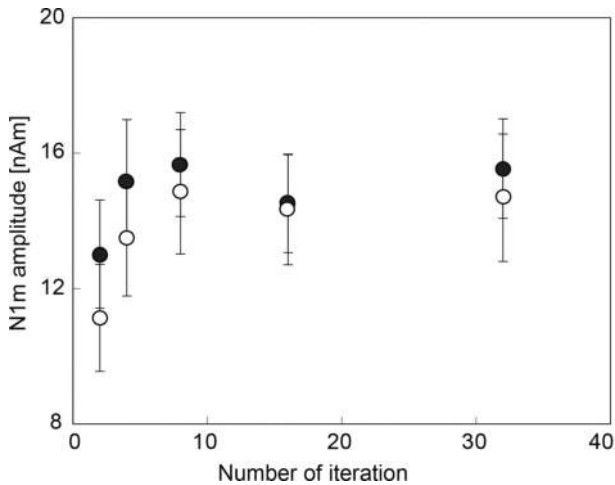

Figure 8. Mean amplitude of the N1m (± standard error) across 10 listeners and hemispheres as a function of the number of iterations with a delay time of 2 ms (○) or 4 ms (●).

Clear N1m responses were observed in both the left and right temporal areas in all listeners as shown in **Figure 7**. The N1m latencies were not systematically affected by the number of iterations of the IRN. **Figure 8** depicts the mean N1m amplitude across 10 listeners as a function of the number of iterations. A greater number of iterations of the IRN, i.e., a larger ϕ_1 value, produced a larger N1m amplitude. This suggests that a stronger pitch produces a larger N1m response. This result is consistent with previous studies [22, 23]. Previously, the amplitude of the AEF component elicited by periodic stimuli was compared with simulated peripheral activity patterns of the auditory nerve [24]. The researchers reported that the amplitude of the N1m was correlated with the pitch strength, estimated on the basis of auditory nerve activity. This finding is consistent with the present results.

Figure 9 shows the relationship between ϕ_1 of the IRN and the N1m amplitude. A larger ϕ_1 value produced a larger N1m response, with a correlation coefficient of 0.76 ($p < 0.05$). However, we found another factor that appears to influence N1m amplitude. To calculate the effects of each ACF factor on AEF responses, we conducted multiple regression analyses with the N1m amplitude as the outcome variable. We used a linear combination of ϕ_1, τ_1 and τ_e as predictive variables in a stepwise fashion. The final version indicated that ϕ_1 and τ_1 were significant factors:

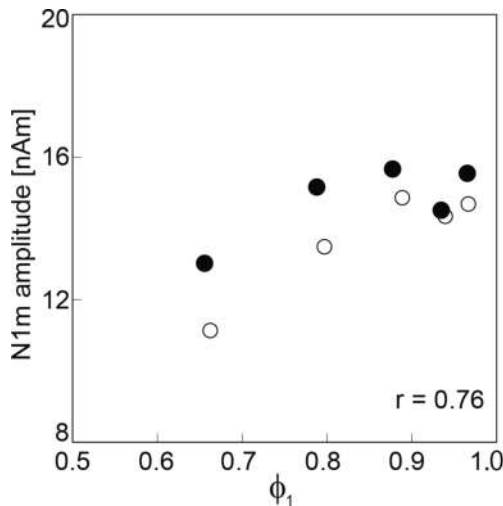$$\text{N1m amplitude} \approx a_1 * \phi_1 + a_2 * \tau_1 + b_1 \tag{3}$$

Figure 9. Relationship between ϕ_1 and mean N1m amplitude. The delay time of the IRN of 2 ms (○) or 4 ms (●).

The model was statistically significant ($p < 0.01$), and the correlation coefficient between the measured and predicted values was 0.88. The standardized partial regression coefficients of the variables a_1 and a_2 in Eq. (3) were 0.77 and 0.44, respectively. These results indicate that

both the ACF factors ϕ_1 and τ_1 had significant effects on N1m responses, although ϕ_1 had a stronger effect.

2.2. AEFs in relation to BPN

To evaluate responses related to ϕ_1 in the auditory cortex, we also recorded the AEFs elicited by BPN with different bandwidths. Eight normal-hearing listeners (22–28 years; all right-handed) took part in the experiment. We produced BPN by repeated digital filtering of 10 s white noise signals. We set the magnitude of the Fourier coefficients to a cut-off slope of 200 dB/octave outside the desired bandwidth. For stimuli with a center frequency of 500 or 1000 Hz, the stimulus bandwidth was set at 1, 40, 80, 160 or 320 Hz. For stimuli with a center frequency of 2000 Hz, the stimulus bandwidth was set at 1, 40, 80, 160, 320 or 640 Hz. The maximum bandwidth was wider than the critical bandwidth for each center frequency [25]. The stimulus duration was 0.5 s, which we took from the 10 s BPN signal and set rise and fall ramps of 10 ms. The sounds were D/A converted with a 16-bit sound card and a sampling rate of 48 kHz. They were presented at a SPL of 74 dB through insert earphones inserted into both the left and right ear canals. **Figure 10** shows the temporal waveforms of the stimuli with a center frequency of 1000 Hz. As the bandwidth of the BPN increases, fluctuations in the envelope of the BPN waveform decrease. The ACF can characterize the BPN, that is, τ_1 corresponds to the center frequency of the BPN and the ϕ_1 value increases as the filter bandwidth decreases.
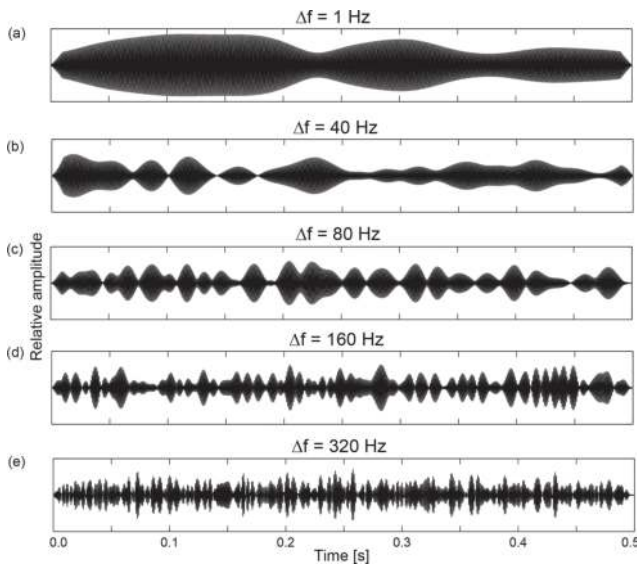

Figure 10. Temporal waveforms of BPNs with a center frequency of 1000 Hz and different bandwidths, Δf, (a) 1 Hz; (b) 40 Hz; (c) 80 Hz; (d) 160 Hz; (e) 320 Hz.

We recorded and analyzed the AEFs using methods similar to previous MEG experiments using IRN. The temporal waveforms of AEFs from 122 channels showed clear N1m responses in both the left and right temporal areas in all listeners. **Figure 11** depicts the mean N1m amplitude across eight listeners as a function of the BPN bandwidths. A narrower BPN bandwidths produced a larger N1m amplitude, that is, the larger the ϕ_1 value, the larger the N1m response. This result is consistent with previous IRN experiments.
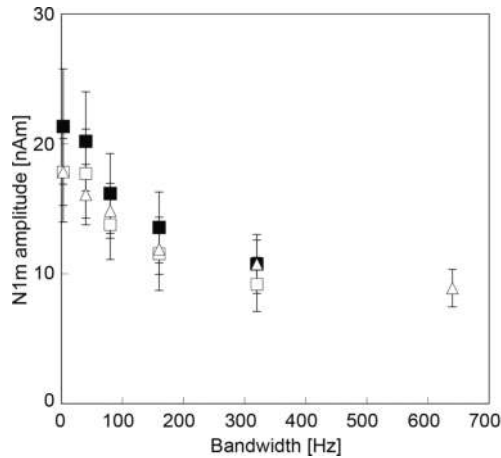

Figure 11. Mean amplitude of the N1m (± standard error) across eight listeners and hemispheres as a function of bandwidth with a center frequency of 500 Hz (□), 1000 Hz (■), and 2000 Hz (△).

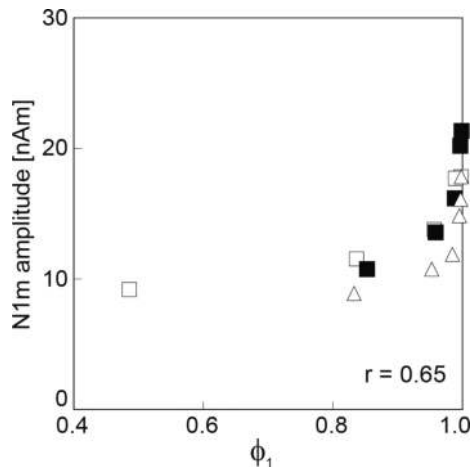**Figure 12.** Relationship between ϕ_1 and mean N1m amplitude. Symbols denote the center frequency of the BPN as 500 Hz (□), 1000 Hz (■), or 2000 Hz (△).

Figure 12 shows the relationship between ϕ_1 of the BPN and the N1m amplitude. A larger ϕ_1 produced a larger N1m response. The correlation coefficient was 0.65 ($p < 0.05$). However, we identified another factor that influences N1m amplitude. To calculate the effects of each ACF factor on AEF response, we conducted multiple regression analyses with the N1m amplitude as the outcome variable. We used a linear combination of ϕ_1, τ_1, and τ_e as predictive variables in a stepwise fashion. The final version indicated that ϕ_1 and τ_e were significant factors:

$$\text{N1m amplitude} \approx a_3 * \phi_1 + a_4 * \tau_e + b_2 \tag{4}$$

The model was statistically significant ($p < 0.01$), and the correlation coefficient between the measured and predicted values was 0.78. The standardized partial regression coefficients of the variables a_3 and a_4 in Eq. (4) were 0.52 and 0.45, respectively. The results indicated that the ACF factors ϕ_1 and τ_e had significant effects on N1m responses.

3. Loudness and annoyance in relation to the effective duration of the ACF, τ_e

3.1. Loudness in relation to IRN

Previous investigations of the relationship between loudness and the BPN bandwidth have concluded that for sounds with the same SPL, loudness remains constant as bandwidth increases, up until the point at which the bandwidth reaches a critical band. For bandwidths larger than the critical band, loudness increases with bandwidth [25]. However, the loudness of a sharply filtered BPN increases with the effective duration of the ACF, i.e., τ_e, even when the bandwidth of the BPN is within the critical band [26]. The τ_e value represents the repetitive components within the signal itself and increases as the BPN bandwidth decreases. However, the envelope and SPL also vary with the BPN bandwidth. This variation of the envelope and SPL might therefore affect the loudness of a BPN signal [27, 28]. To eliminate the effects of these factors, we investigated the effects of τ_e on loudness using IRN. The envelope and SPL variation of the IRN are much smaller than those of the BPN [29].

We produced IRN by applying a delay-and-add algorithm to the BPN that was filtered from white noise using the fourth-order Butterworth filters ranging between 100 and 3500 Hz. The number of iterations of the delay-and-add process was set at 2, 4, 8, 16, and 32. The delay values were set at 0.5, 1, 2, 4, 8, and 16 ms, corresponding to pitches of 2000, 1000, 500, 250, 125, and 62.5 Hz, respectively. The duration of the stimuli was 0.5 s and the rise and fall ramps were 10 ms. The sounds were D/A converted with a 16-bit sound card and sampling rate of 48 kHz. The sounds were presented at a SPL of 60 dB through insert earphones inserted into the left and right ear canals. **Figure 13** shows the τ_e and ϕ_1 values of the IRN used in the experiment.
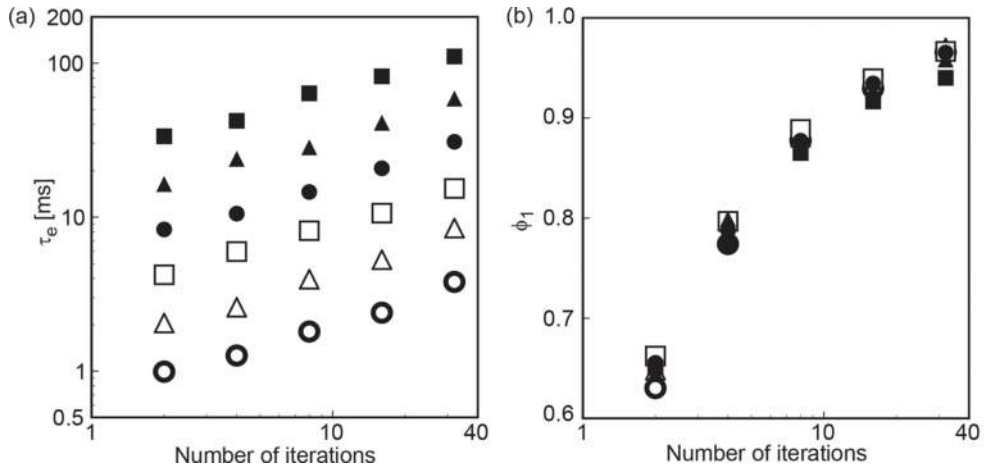
Figure 13. (a) τ_e and (b) ϕ_1 of the IRN used in the experiment as a function of the number of iterations with delays of (○) 0.5, (△) 1, (□) 2, (●) 4, (▲) 8 and (■) 16 ms.

Ten listeners (aged 21–37 years) with normal hearing took part in the experiment. We obtained loudness matches using a two-interval, adaptive forced-choice procedure converging on the point of subjective equality (PSE) following a simple 1-up, 1-down rule [30]. The experiment took place in a soundproof room. In each trial, the fixed (test) and variable (reference) sounds were presented in randomized order with equal probability at an interval of 500 ms. The test sound was an IRN and the reference sound was a 1-kHz pure tone. The listener was asked to indicate which sound they perceived as louder by pressing a key on a keyboard. For each
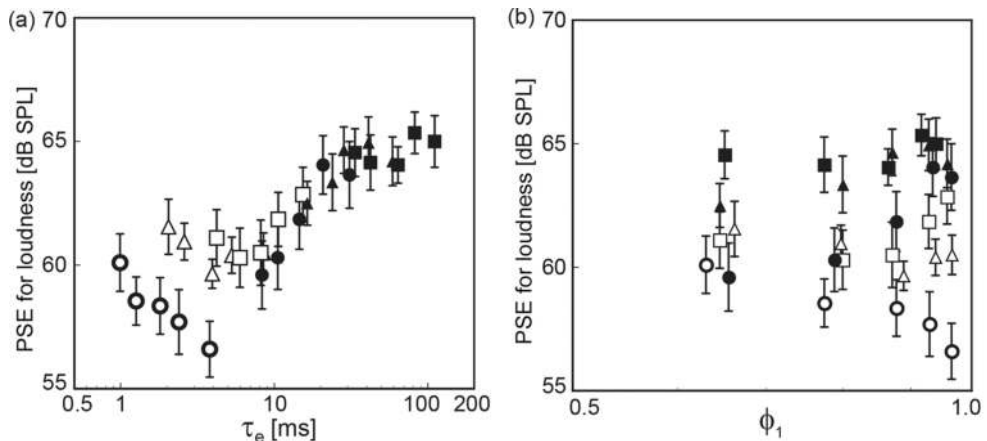

Figure 14. Mean PSE for loudness (± standard error) across 10 listeners as a function of (a) τ_e and (b) ϕ_1 for IRN with a delay of (△) 0.5, (□) 1, (○) 2, (●) 4, (▲) 8 and (■) 16 ms.

adaptive track, the overall level of the test sound was fixed at 60 dB SPL, and the starting level of the reference sound was 50 dB SPL. The level of the reference sound was controlled with an adaptive procedure: when the listener judged the reference sound to be louder than the test sound, the SPL of the test sound was lowered by a given amount, and when the listener judged the test sound to be louder than the reference sound, the SPL of the reference sound was increased by that same amount.

Figure 14 shows the PSE for loudness as a function of τ_e and ϕ_1 of the IRN. ϕ_1 was not correlated with the perceived loudness. When τ_e was between 10 and 100 ms, the perceived loudness increased with τ_e, clearly confirming that loudness is influenced by the repetitive components of sounds [26] in the τ_e range between 10 and 100 ms. The increase in loudness for the τ_e values between 10 and 100 ms was approximately 5 dB.

When τ_e was less than 5 ms, the loudness of the IRN increased with decreasing τ_e and the bandwidth of the IRN was larger than the critical bandwidth. These tendencies may explain the basis of the critical band effect, such that loudness remains constant as the bandwidth of the noise is narrower than the critical band, then increases with increasing bandwidth beyond the critical band [25]. Loudness models are able to predict these tendencies [31, 32].

The loudness model introduced previously [31, 32] was unable to predict loudness when the delays were 2 and 4 ms for stimuli with a pitch of 500 and 250 Hz, respectively. Loudness increases caused by a tonal component are predictable according to τ_e in a certain range. Previous studies have indicated that the τ_e values of various noise sources, such as airplanes [33], trains [34], motor bikes [35] and flushing toilets [36], are within the range of 1–200 ms. This suggests that τ_e is a useful criteria for measuring the loudness of various sounds. Thus, this value is likely helpful for the identification of sound sources.

3.2. Annoyance in relation to BPN

Annoyance is one of the most commonly studied features of environmental noise [37]. Basically, psychoacoustic annoyance depends on loudness and other factors such as timbre and the temporal structure of sounds. Loudness and annoyance have been distinguished previously: Annoyance is the reaction of an individual to noise within the context of a given situation, while loudness is directly related to SPL [38]. To evaluate whether annoyance is related to the effective duration of the ACF, i.e., τ_e, we examined the annoyance elicited by a pure tone and BPN stimuli with different bandwidths.

We used pure tone and BPN signals with center frequencies of 1000 and 2000 Hz as auditory signals. We used a maximum length bandpass filtered sequence signal (order 21; sampling frequency, 44,100 Hz) as the basic stimulus. To control the ACF of the BPN, we varied the filter bandwidth at 0, 40, 80, 160, and 320 Hz using a cut-off slope of 2068 dB/octave. The sounds were D/A converted with a 16-bit sound card and sampling rate of 48 kHz. The sounds were presented to both the left and right ears at an SPL of 74 dBA using headphones (Sennheiser HD-340). **Figure 15** shows τ_e of the stimuli used in the experiment.
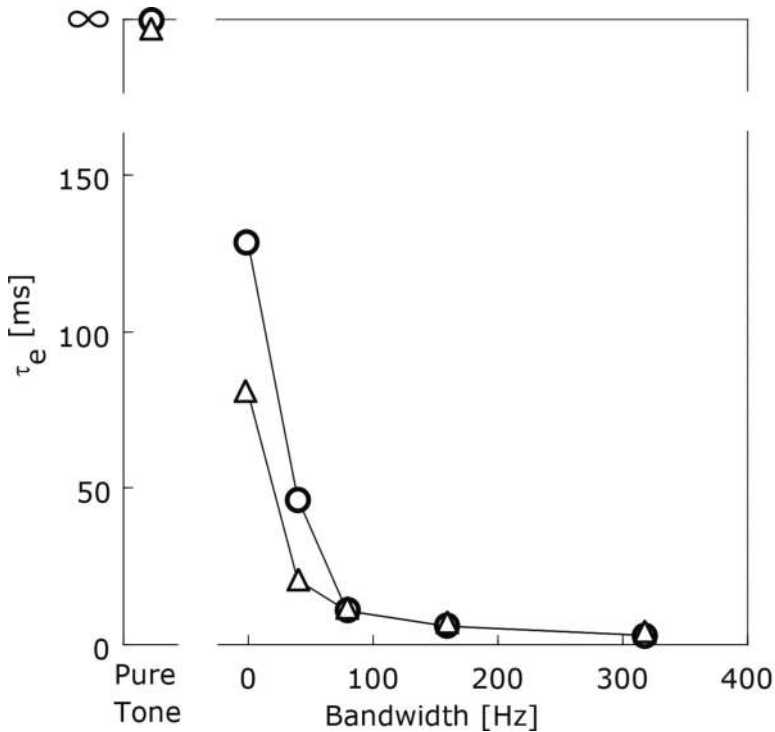
Figure 15. The measured effective duration of NACF, i.e., τ_e, of the signal as a function of the bandwidth. Different symbols indicate different frequencies: (◯): 1000 Hz; (△): 2000 Hz.

Eight listeners aged 21–23 years with normal hearing took part in the experiment. We performed paired-comparison tests for all combinations of the pairs of the pure tone and BPN stimuli. The duration of the stimuli was 2.0 s, the rise and fall times were 50 ms, the silent interval between the stimuli was 1.0 s, and the interval between the pairs was 3.0 s, which was the time during which the listeners were expected to make a response. They were asked to judge which of the two sound signals was more annoying. We calculated the scale values of the annoyance rated by each listener according to Case V of Thurstone's theory [39].

The relationship between the scale values of annoyance and τ_e is shown in **Figure 16**. The averaged scale values of annoyance increased as τ_e increased within the critical band for both center frequencies of 1000 and 2000 Hz. The τ_e value represents the repetitive feature or tonal component of the auditory signals. Previous research suggests that tonal components increase the perceived annoyance and noisiness of broadband noise [35, 40, 41]. This is consistent with the present results. Two of the eight listeners reported the least annoyance for pure tone stimuli, with BPN stimuli with the widest bandwidth and a center frequency of 2000 Hz rated as the most annoying. In other words, annoyance increased as τ_e decreased. This could indicate that the effects of τ_e on annoyance are subject to individual variation.
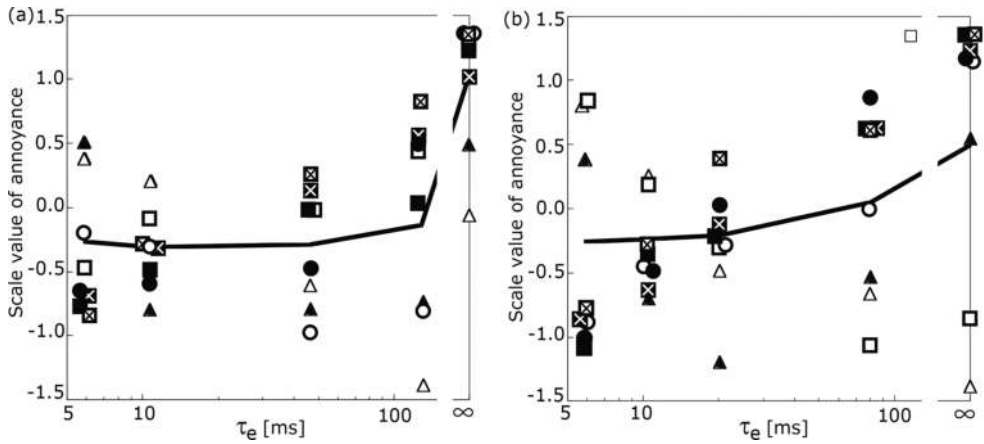
Figure 16. Scale value of annoyance as a function of τ_e for BPN with a center frequency of (a) 1000 Hz and (b) 2000 Hz. Each symbol represents one listener. The line represents the mean scale value of the eight listeners.

4. Concluding remarks

In this study, we investigate the effects of ACF factors on physiological and psychological responses. As a result, we found that the ACF factors ϕ_1, τ_1, and τ_e had significant effects on N1m response, suggesting that ACF factors are used as cues in the auditory cortex. We also found that the ACF factors ϕ_1 and τ_e influence loudness and annoyance, suggesting that ACF factors are used as a cue for perception. These results indicate that the human auditory system has an autocorrelation-like mechanism.

Acknowledgements

Appendix

Auditory evoked fields (AEFs): Magnetic fields evoked by any abrupt sound or change in a continuous sound in the auditory cortex.

Butterworth filters: A kind of signal processing filter widely used for bandpass filtering.

Bandpass filtered noise (BPN): A noise in which frequency components are limited by bandpass filtering.

Envelope: Approximate shape of a sound wave form calculated by joining the peak amplitude.

Equivalent current dipole: A dipole estimated from the measured magnetic fields in the human brain, widely used in MEG analysis.

Interspike interval: Observed time between spikes from a single neuron.

Magnetoencephalography (MEG): A noninvasive technique for investigating human brain activity by measuring the magnetic fields produced by electric currents flowing in neurons.

Multiple regression analysis: A statistical technique for predicting a dependent variable from independent variables.

Paired-comparison tests: A psychophysical method that measures a linear distance among paired stimuli.

Point of subjective equality (PSE): Any of the points along a stimulus dimension at which a variable stimulus is judged by a listener to be equal to a standard stimulus.

Pure tone: A tone with a frequency component and a sinusoidal wave.

Tonotopic organization: Spatial arrangements in which sounds of different frequencies are processed in the auditory system.

Two-interval, adaptive forced-choice procedure: A psychophysical experimental design in which listeners are instructed to make a response between two alternatives within a timed interval, and the next alternative is determined by the previous response.

White noise: A random signal with equal intensity at all frequencies.

Author details

Yoshiharu Soeta

Address all correspondence to: y.soeta@aist.go.jp

National Institute of Advanced Industrial Science and Technology (AIST), Osaka, Japan

References

[1] Soeta Y, Ando Y. Neurally based measurement and evaluation of environmental noise. Tokyo: Springer Japan; 2015. DOI: 10.1007/978-4-431-55432-5.

[2] Yost WA. Pitch strength of iterated ripple noise. Journal of the Acoustical Society of America 1996;100:3329–3335. DOI: 10.1121/1.416973.

[3]   Ando Y. Auditory and visual sensations. New York: Springer; 2010. DOI: 10.1007/b13253.

[4]   Ando Y. Architectural acoustics: blending sound sources, sound fields, and listeners. New York: Springer-Verlag; 1998.

[5]   Licklider JCR. A duplex theory of pitch perception. Experimenta. 1951;7:128–134. DOI: 10.1007/BF02156143.

[6]   Wightman FL. The pattern-transformation model of pitch. Journal of the Acoustical Society of America 1973;54:407–416. DOI: 10.1121/1.1913592.

[7]   Meddis R, Hewitt M. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. Journal of the Acoustical Society of America. 1991;89:2866–2882. DOI: 10.1121/1.400725.

[8]   Fay RR, Yost WA, Coombs S. Psychophysics and neurophysiology of repetition noise processing in a vertebrate auditory system. Hearing Research 1983;12:31–55. DOI: 10.1016/0378-5955(83)90117-X.

[9]   ten Kate JH, van Bekkum MF. Synchrony-dependent autocorrelation in eighth-nerve-fiber response to rippled noise. Journal of the Acoustical Society of America 1988;84:2092–2102. DOI: 10.1121/1.397054.

[10] Shofner WP. Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. Journal of the Acoustical Society of America 1991;90:2450–2466. DOI: 10.1121/1.402049.

[11] Shofner WP. Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural autocorrelograms. Journal of Neurophysiology 1999;81:2662–2674.

[12] Winter IM, Wiegrebe L, Patterson RD. The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. Journal of Physiology 2001;537:553–566. DOI: 10.1111/j.1469-7793.2001.00553.x.

[13] Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. Journal of Neurophysiology. 1996;76:1698–1716.

[14] Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. Journal of Neurophysiology. 1996;76:1717–1734.

[15] Hämäläinen MS, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 1993;65:413–497. DOI: 10.1103/RevModPhys.65.413.

[16] Elberling C, Bak C, Kofoed B, Lebech J, Sarmark G. Auditory magnetic fields from the human cerebral cortex: location and strength of an equivalent current dipole. Acta Neurologica Scandinavica 1982;65:553–569. DOI: 10.1111/j.1600-0404.1982.tb03110.x.

[17] Romani GL, Williamson SJ, Kaufman L. Tonotopic organization of the human auditory cortex. Science 1982;216:1339–1340. DOI: 10.1126/science.7079770.

[18] Pantev C, Hoke M, Lehnertz K, Lütkenhöner B, Anogianakis G, Wittkowski W. Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalography and Clinical Neurophysiology 1988;69:160–170. DOI: 10.1016/0013-4694(88)90211-8.

[19] Langner G, Sams M, Heli P, Schulze H. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: evidence from magnetoencephalography. Journal of Comparative Physiology A 1997;181:665–676. DOI: 10.1007/s003590050148.

[20] Cansino S, Ducorps A, Ragot R. Tonotopic cortical representation of periodic complex sounds. Human Brain Mapping 2003;20:71–81. DOI: 10.1002/hbm.10132.

[21] Näätänen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 1987;24:375–425. DOI: 10.1111/j.1469-8986.1987.tb00311.x.

[22] Soeta Y, Nakagawa S, Tonoike M. Auditory evoked magnetic fields in relation to the iterated rippled noise. Hearing Research 2005;205:256–261. DOI: 10.1016/j.heares.2005.03.026.

[23] Krumbholz K, Patterson RD, Seither-Preisler A, Lammertmann C, Lütkenhöner B. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cerebral Cortex 2003;13:765–772. DOI: 10.1093/cercor/13.7.765.

[24] Seither-Preisler A, Krumbholz K, Lutkenhoner B. Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? Audiology Neurootology. 2003;8:322–337. DOI: 10.1159/000073517.

[25] Zwicker E, Flottorp G, Stevens SS. Critical bandwidth in loudness summation. Journal of the Acoustical Society of America 1957;29:548–557. DOI: 10.1121/1.1908963.

[26] Sato S, Kitamura T, Ando Y. Loudness of sharply (2068 dB/Octave) filtered noises in relation to the factors extracted from the autocorrelation function. Journal of Sound and Vibration 2002;250:47–52. DOI: 10.1006/jsvi.2001.3888.

[27] Zhang C, Zeng FG. Loudness of dynamic stimuli in acoustic and electric hearing. Journal of the Acoustical Society of America 1997;102:2925–2934. DOI: 10.1121/1.420347.

[28] Moore BCJ, Vickers D, Baer T, Launer S. Factors affecting the loudness of modulated sounds. Journal of the Acoustical Society of America 1999;105:2757–2772. DOI: 10.1121/1.426893.

[29] Soeta Y, Nakagawa S. Effect of the repetitive components of a noise on loudness. Journal of Temporal Design in Architecture and the Environment. 2008;8:1–7.

[30] Levitt H. Transformed up–down procedures in psychophysics. Journal of the Acoustical Society of America 1971;49:467–477. DOI: 10.1121/1.1912375.

[31] Moore BCJ, Glasberg BR, Baer T. A model for the prediction of thresholds, loudness, and partial loudness. Journal of the Audio Engineering Society 1997;45:224–240.

[32] Zwicker E, Fastl H. Psychoacoustics. Facts and models. New York: Springer; 2010. 1999. DOI: 10.1007/978-3-662-09562-1.

[33] Fujii K, Soeta Y, Ando Y. Acoustical properties of aircraft noise measured by temporal and spatial factors. Journal of Sound and Vibration 2001;241:69–78. DOI: 10.1006/jsvi.2000.3278.

[34] Sakai H, Hotehama T, Prodi N, Pompoli R, Ando Y. Diagnostic system based on the human auditory-brain model for measuring environmental noise – an application to the railway noise. Journal of Sound and Vibration 2002;250:9–21. DOI: 10.1006/jsvi.2001.3884.

[35] Fujii K, Atagi J, Ando Y. Temporal and spatial factors of traffic noise and its annoyance. Journal of Temporal Design in Architecture and the Environment. 2002;2:33–41.

[36] Kitamura T, Shimokura R, Sato S, Ando Y. Measurement of temporal and spatial factors of a flushing toilet noise in a downstairs bedroom. Journal of Temporal Design in Architecture and the Environment. 2002;2:13–19.

[37] Berglund B, Berglund U, Lindvall T. Scaling loudness, noisiness, and annoyance of aircraft noise. Journal of the Acoustical Society of America 1975;57:930–934. DOI: 10.1121/1.380535.

[38] Hellman RP. Loudness, annoyance, and noisiness produced by single-tone-noise complexes. Journal of the Acoustical Society of America 1982;72:62–73. DOI: 10.1121/1.388025.

[39] Thurstone LL. A law of comparative judgment. Psychological Review 1927;34:273–289.

[40] Kryter KD, Pearsons KS. Judged noisiness of a band of random noise containing an audible pure tone. Journal of the Acoustical Society of America 1965;38:106–112. DOI: 10.1121/1.1909578.

[41] Hargest TJ, Pinker RA. The influence of added narrow band noises and tones on the subjective response to shaped white noise. Journal of the Royal Aeronautical Society. 1967;71:428–430. DOI: 10.1017/S0001924000055512.