Forum Acusticum 2011 - Aalborg - Denmark

Contribution of the early part of the head-related impulse response to the formation of two spectral notches of vertical localization cues

Kazuhiro Iida, Naokazu Gamoh, Yohji Ishii
Faculty of Engineering, Chiba Institute of Technology, Narashino, Japan.
Masayuki Morimoto
Graduate School of Engineering, Kobe University, Kobe, Japan.

Summary
The authors have previously shown that the first and the second spectral notches (N1 and N2) above 4 kHz in the head-related transfer function play an important role as spectral cues for vertical localization. In the present study, the authors analyzed which part of the head-related impulse response (HRIR) forms N1 and N2, using front-direction HRIRs measured at sampling frequencies of 48 and 96 kHz. The results of the analysis show that (1) both N1 and N2 are formed by the very early part of the HRIR (within 0.5ms); (2) a unique response pattern is observed in the early part of the HRIR; (3) N1 and N2 are not formed by simple interference between a direct wave and a single reflection.

PACS no. 43.66.Qp, 43.66.Pn

1. Introduction

It is well known that the spectral information contained in head-related transfer functions (HRTFs) is a cue for vertical localization. Iida *et al.* [1] proposed a parametric HRTF model to clarify the contribution of each spectral peak and notch as a spectral cue for vertical localization. The parametric HRTF is recomposed using all or some of the spectral peaks and notches above 4 kHz extracted from the measured HRTF, which are expressed parametrically by their frequency, level, and sharpness. Based on the results of sound localization tests, they found that a parametric HRTF recomposed using the first and second notches (N1 and N2) and the first peak (P1) provided almost the same localization accuracy as the measured HRTFs. In addition, the frequency of N1 and N2 was found to depend strongly on the source elevation, whereas that of P1 was independent of elevation. They concluded that N1 and N2 can be regarded as spectral cues, and that the hearing system could utilize P1 as the reference information to analyze N1 and N2.

Takemoto *et al.* [2] calculated HRTFs and pinna-related transfer functions (PRTFs) in the median plane from head shape data measured by magnetic resonance imaging using a finite-difference time-domain method. The results showed that N1, N2 and P1 were observed in PRTFs as well as in HRTFs. However, the mechanism for the formation of N1 and N2 is still unclear.

This study analyzes which part of head-related impulse responses (HRIRs) contributes to the formation of N1 and N2 using a rectangular time window function, which clips the early part of HRIRs at zero-crossing samples. Furthermore, the mechanism for the formation of N1 and N2 is discussed.

2. Analysis of early part of HRIRs using rectangular time window function

2.1. Algorithm for spectral analysis of HRIRs

The front-direction HRIRs of four subjects (IST, ISY, IT, and ID) were analyzed at a sampling frequency of 48 kHz for 512 samples. The algorithm used to calculate the amplitude spectrum is as follows:

Step 1: find the maximum HRIR absolute amplitude.

Step 2: clip the HRIR using a rectangular time window function whose origin is the 1st of the 512 samples and whose end point is the Nth sample

from the maximum obtained in step 1, where N is a zero-crossing sample (Figure 1).
Step 3: pad the clipped HRIR from its end point to the 512th sample with zero values.
Step 4: obtain the amplitude spectrum of the resulting 512 sample data by FFT.

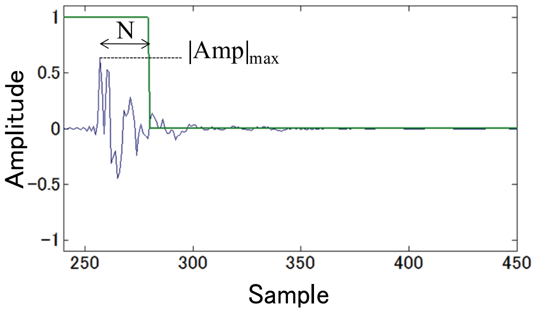

Figure 1. HRIR and rectangular window function.

2.2. Results of analysis

The calculated HRIR spectrum for the left ear of subject IST is shown in Figure 2, superimposed upon the full-length HRIR. For N=5, a single spectral notch is observed. However, its frequency does not correspond to that of either N1 or N2. For N=11, both N1 and P1 appear. N1, N2, and P1 are all present for N=16, although, N2 is not very deep. For N=23, N1, N2, and N3 are observed and N2 is as deep as that in the full-length HRIR. These tendencies are commonly observed for all subjects.
Figure 3 shows the minimum value of N for which the frequencies of the two spectral notches match those of N1 and N2. For all subjects, N1 and N2 appeared for the range N=20-23.
These results show that the early part (approximately 0.5 ms) of the HRIR plays an important role in the formation of N1 and N2.

3. Discussion

3.1. Unique pattern in the early part of HRIRs

Figure 4 shows the waveforms of the early part (N=20-23) of HRIRs for the four subjects. The red and blue dotted lines denote the amplitude spectrum of the early part of the HRIR and the full-length HRIR, respectively. A distinctive pattern of peaks and valleys, “PPVVPPVV”, can be seen in the waveforms.
To examine this response pattern at higher resolution, HRIRs were measured using a sampling frequency of 96 kHz. Figure 5 shows the front-direction HRIRs for subject IST obtained using sampling frequencies of 96 and 48 kHz. It can be seen that the same unique response pattern appears at both sampling frequencies.
These results infer that for the front direction, N1 and N2 are formed by the “PPVVPPVV” response pattern.

3.2. Formation mechanism for N1 and N2

The relation between the “PPVVPPVV” response pattern and the formation mechanism for N1 and N2 will now be considered.
Raykar *et al.* [3] assumed that spectral notches are generated by interference of a direct wave and a reflection from the concha wall, and suggested that notch frequencies can be estimated using equation (1).

$$f_n(\varphi) = \frac{(2n+1)}{2t_d(\varphi)} \text{ , } n = 0,1,\cdots \qquad (1)$$

Here, f_n is the notch frequency [Hz], t_d is the arrival time difference between the direct wave and the reflection [s], φ is the incident angle of the direct wave, and n is the notch number.
Using equation (1), f_0 and f_1 were calculated assuming that the 1st P in “PPVVPPVV” corresponds to the direct wave and the 2nd P corresponds to the reflection. Table I shows the calculated values of f_0 and f_1 and the frequencies of N1 and N2 for the four subjects. Except for N1 for subject IT, there is no agreement between f_0 and N1 or f_1 and N2. These results indicate that N1 and N2 are not formed by simple interference between a direct wave and a single reflection.
On the other hand, Takemoto *et al.* [2] noted that resonance in the upper cavities (cymba conchae, triangular fossa, and scaphoid fossa) affects the amplitude and frequency of spectral notches. Furthermore, they showed that an incident wave from the upper cavities with a time delay of approximately 0.2 ms plays an important role in forming the spectral notches [4]. This incident wave corresponds to the 3rd or 4th P in the “PPVVPPVV” response.
Figure 6 shows that the period of the “PPVV” is approximately 0.25 ms, which suggests that it may be responsible for P1 (4kHz). Note that P1 is considered to be resonance in the concha [5].

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

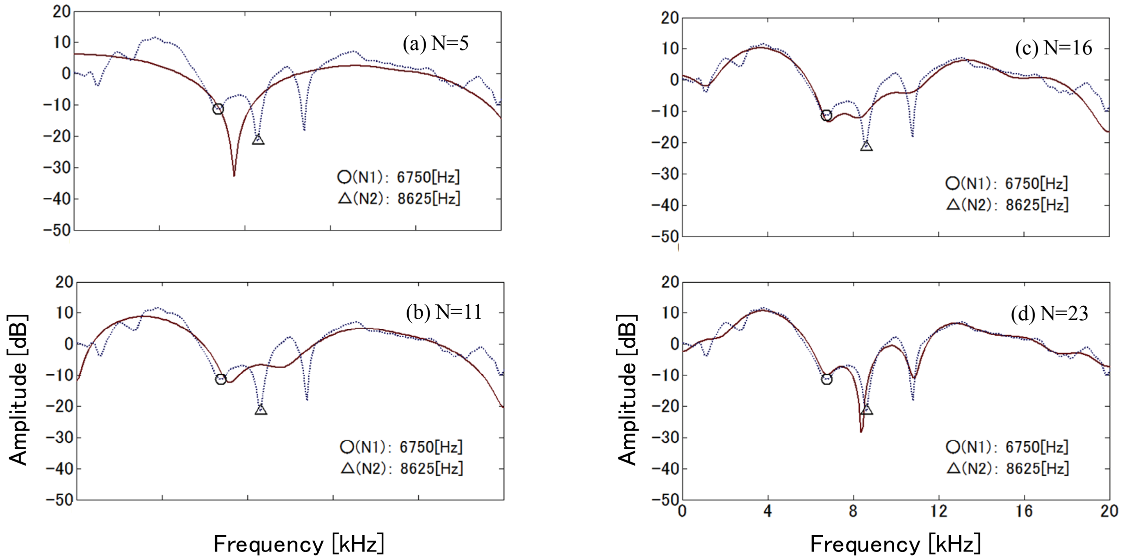

Figure 2. Amplitude spectrum of front-direction HRIR for subject IST at the left ear. Red solid lines: various N values (see Figure 1), blue dotted lines: N=256 (full-length).

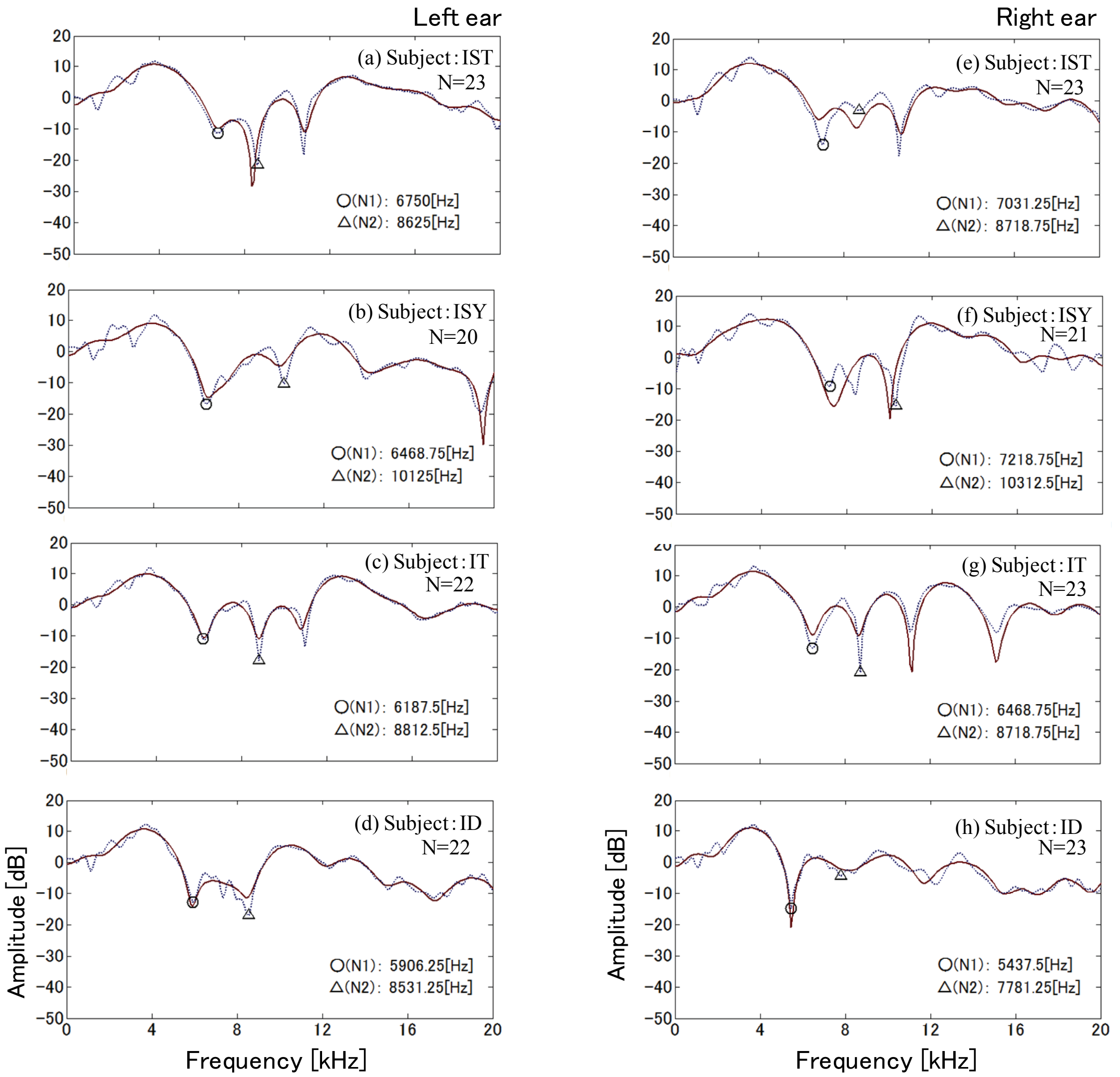

Figure 3. Amplitude spectrum of front-direction HRIR for four subjects. Red solid lines: N=20-23 (approximately 0.5ms), blue dotted lines: N=256 (full-length).

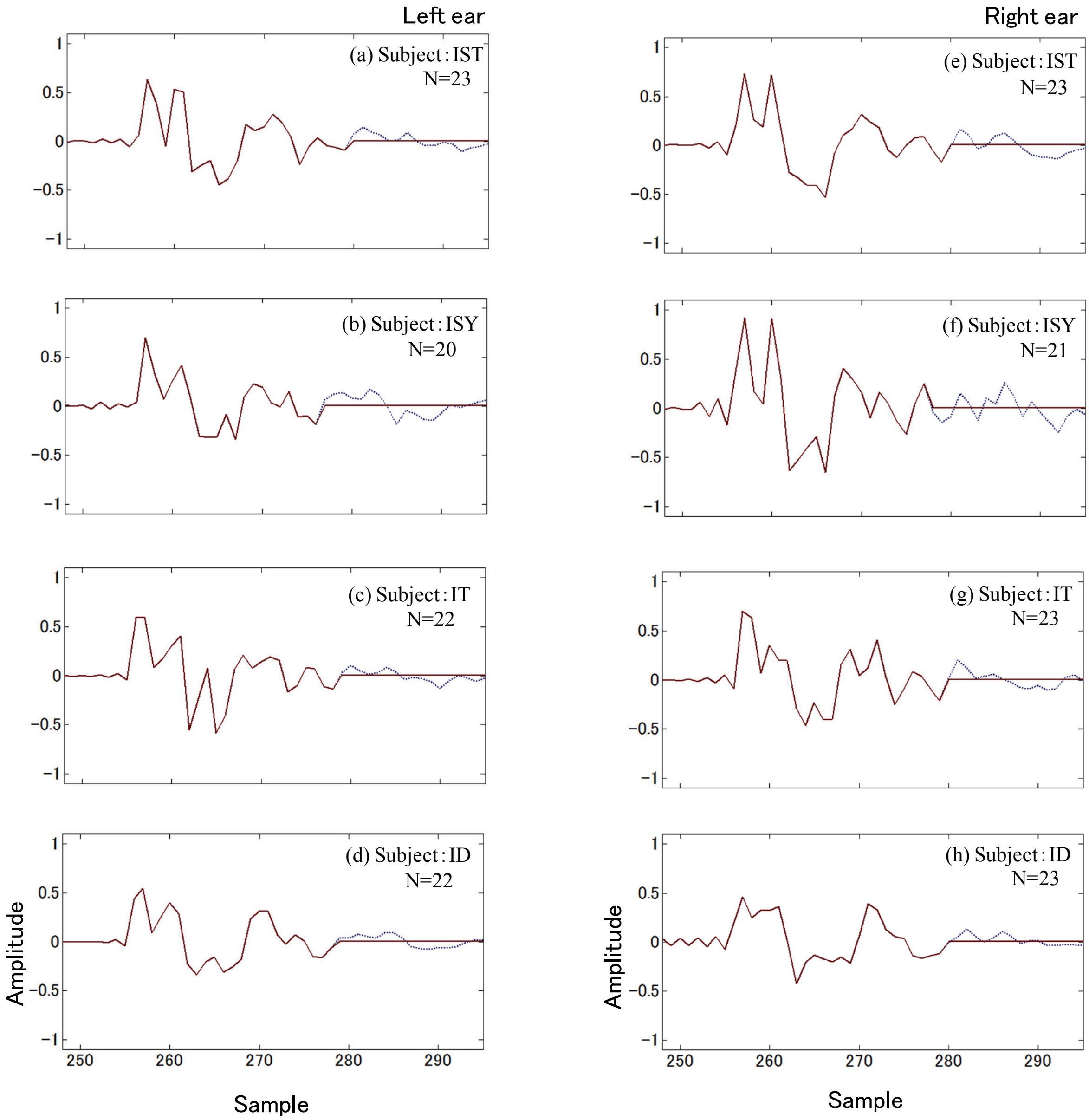

Figure 4. Front-direction HRIR. Red solid lines: extracted response using rectangular window function with length N of 20-23, blue dotted lines: original response.

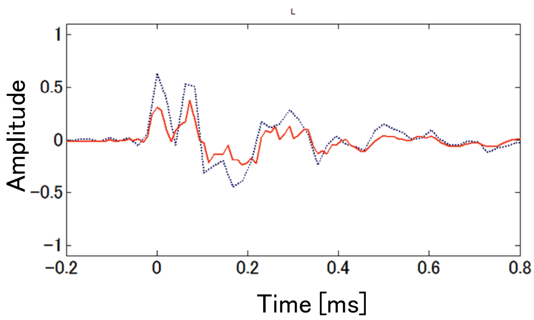

Figure 5. Front-direction HRIR for subject IST at the left ear for sampling frequencies of 96 kHz (red solid line) and 48 kHz (blue dotted line).

Table I. Comparison between calculated f_0 and f_1 and measured N1 and N2 frequencies.

Subject	IST	ISY	IT	ID
t_d [ms]	0.0625	0.0833	0.0833	0.0625
f_0 [Hz]	8000	6000	6000	8000
f_1 [Hz]	24000	18000	18000	24000
N1 [Hz]	6750	6468.75	6187.5	5906.25
N2 [Hz]	8625	10125	8812.5	8531.25

(c) European Acoustics Association, ISBN: 978-84-694-1520-7, ISSN: 221-3767

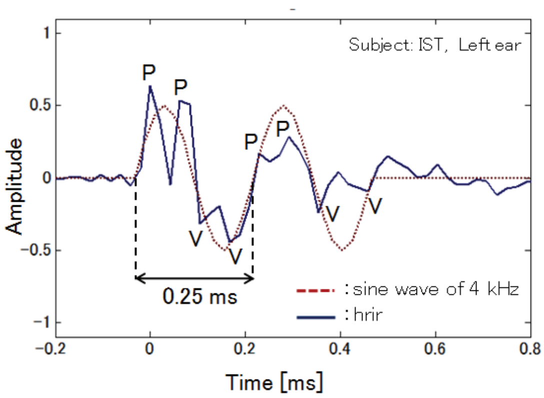

Figure 6. Early part of HRIR and sine wave with frequency of 4 kHz.

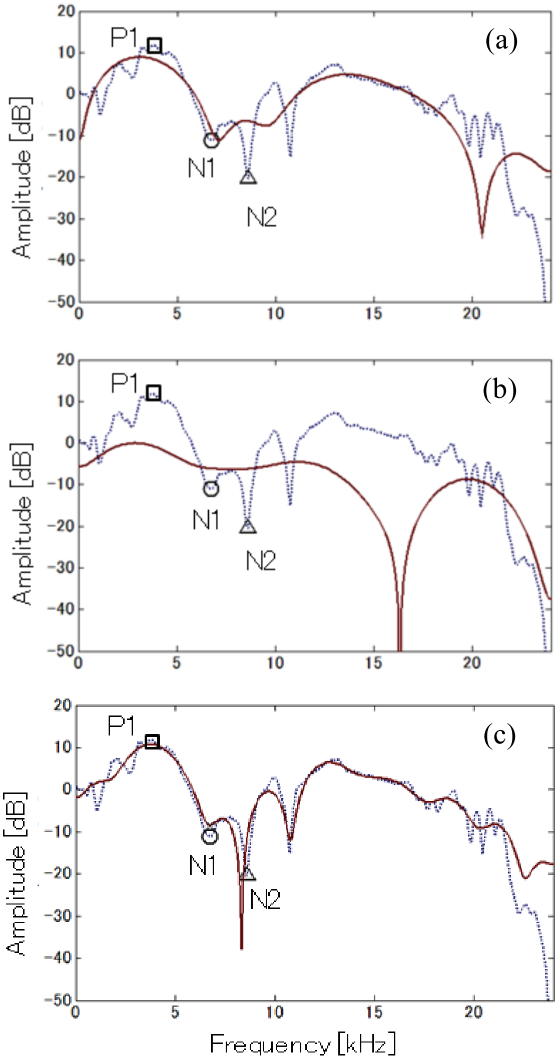

Figure 7. Amplitude spectrum of 1st “PPVV” (a), 2nd “PPVV” (b), and entire “PPVVPPVV” (c).

Next, the amplitude spectrum of the 1st “PPVV” pattern, the 2nd “PPVV” pattern, and the entire “PPVVPPVV” response was analyzed, and the results are shown in Figure 7(a), (b), and (c), respectively. It can be seen that although N1 and P1 appear in Figure 7(a), neither N1 nor N2 appear in Figure 7(b). Since N2 is not found in either the 1st or 2nd “PPVV” pattern, it must be produced by the full sequence “PPVVPPVV” (Figure 7 (c)).

4. Conclusion

This study analyzed which part of HRIRs contribute to the formation of N1 and N2 using a rectangular time window function, which clips the early part of HRIRs at zero-crossing samples. The results of the analysis show that (1) both N1 and N2 are formed by the very early part of the HRIR (within 0.5ms); (2) a unique response pattern is observed in the early part of the HRIR; (3) N1 and N2 are not formed by simple interference between a direct wave and a single reflection.

Acknowledgement

This work was supported in part by Grant-in-Aid for Scientific Research (A) 22241040.

References

[1] K. Iida *et al.*: Median plane localization using parametric model of the head-related transfer function based on spectral cues. Applied Acoustics 68 (2007) 835-850.

[2] H. Takemoto *et al.*: A simple pinna model for generating head-related transfer functions in the median plane. Proc. 20th International Congress on Acoustics (2010).

[3] V. C. Rakar *et al.*: Extracting the frequencies of the pinna spectral notches in measured head related impulse responses.J.A.S.A. 118 (2005) 364-374.

[4] H. Takemoto *et al.*: Analysis of early-reflected sound wave at the pinna. Proc. Spring Meeting of Acoustical Society of Japan (2011.03). (in Japanese)

[5] E. A. G. Shaw, R, Teranishi: Sound pressure generated in an external-ear replica and real human ears by a nearby point source. J. Acoust. Soc. Am. 44 (1968) 240-249.