Lateralization of Incoherent Narrow-Band Noise Bursts

Juha Merimaa[1], Christof Faller[2]

[1] *Lab. of Acoustics and Audio Signal Processing, Helsinki University of Technology, Finland, Email: juha.merimaa@tkk.fi*

[2] *Audiovisual Communications Laboratory, EPFL Lausanne, Switzerland, Email: christof.faller@epfl.ch*

Introduction

Faller and Merimaa [1] proposed an auditory localization model where the contribution of the instantaneous (as evaluated by the auditory system) interaural time differences (ITDs) and interaural level differences (ILDs) to the localization is controlled by the instantaneous interaural coherence (IC). In the most simple form, the model was implemented such that if the IC was above a set threshold, the ITD and ILD cues were selected as to contribute to the localization judgement and otherwise discarded. The purpose of the subjective experiment reported in this paper was to directly study the effect of coherence on localization.

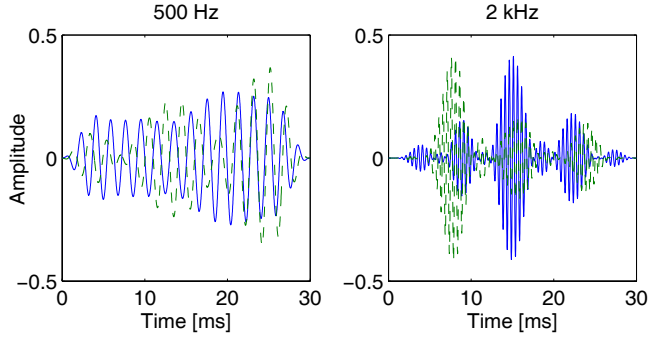

Figure 1: Examples of left and right samples with a center frequency of 500 Hz (left panel) and 2 kHz (right panel).

Stimuli and procedure

Since full incoherence is difficult to create in free-field, the study was conducted as a lateralization experiment using headphone reproduction. The lateralization was investigated at two frequency bands: At the lower tested band, the processing of the ITDs and IC is based on waveform features and at the higher band on envelopes.

The duration of all stimuli was 30 ms and they were gated on and off with 10 ms raised cosine ramps. The low-frequency stimuli consisted of 15 Gaussian noise samples bandpass filtered between 450–550 Hz. In each case, independent noise samples were used for the left and right ear signals and it was verified that the resulting waveform coherence of each sample pair was below 0.05. The 15 high-frequency stimuli were so called transposed stimuli [2]. Initially, 15 pairs of Gaussian noise samples lowpass-filtered at 200 Hz and with a waveform coherence below 0.05 were created. The samples were subsequently half-wave rectified, lowpass filtered at 200 Hz again, and used as envelopes modulating 2 kHz sinusoidal carrier signals. This resulted in signal pairs with energy between 1800–2200 Hz and incoherent envelopes. The stimuli are illustrated in Figure 1. The overall ILD of each stimulus pair was zero, although the instantaneous ILDs varied considerably.

In addition to the incoherent stimuli, 5 fully coherent anchors and a pointer with similar frequency characteristics as the stimuli were created for both frequency bands. The anchors had a zero ILD and constant ITDs. The ITD of the pointer was always zero.

The task of the listeners was to adjust the ILD of the pointer to match the dominant lateralization of each stimulus using a slider on a graphical user interface (GUI). For each stimulus, there was also a possibility to report perception of two separate auditory events by using a corresponding check box on the GUI. Within a single evaluation, the current stimulus was repeated at a rate of 4 Hz and switched to a corresponding repetition of the pointer while the subject was holding down the mouse button on the slider. The listener was free to use as much time for the evaluation as needed. Each listening session consisted of a single evaluation of each stimulus at one frequency band in a randomized order. Six listeners participated in the experiment. Each listener completed 8 sessions, thus evaluating each stimulus four times.

Results

The results of the experiment are shown in Figure 2. The coherent anchors were displaced from the center according to the applied ITDs, as expected. All incoherent stimuli also resulted in a consistently localizable auditory event. However, the lateralization varied considerably between the stimuli. The variances of the lateralization judgements of the incoherent stimuli at the 500 Hz center frequency were slightly higher than those of the anchors but, on average, did not differ at 2 kHz. In some cases, there were also individual differences in the perceived lateralization. For the incoherent stimuli, the subjects reported perception of two distinct auditory events in 20–50% of the evaluations. This fairly frequent perception of multiple auditory events might explain the larger variations and individual differences.

Modeling

The stimuli were studied further with the auditory model described in [1]. The peripheral part of the model consisted of a Gammatone filterbank, half-wave rectification, envelope compression, and lowpass filtering. The instantaneous binaural cues were analyzed as a function of time, averaged over a sliding exponentially decaying window with a time constant of 10 ms. The analysis was performed at several overlapping critical bands to take

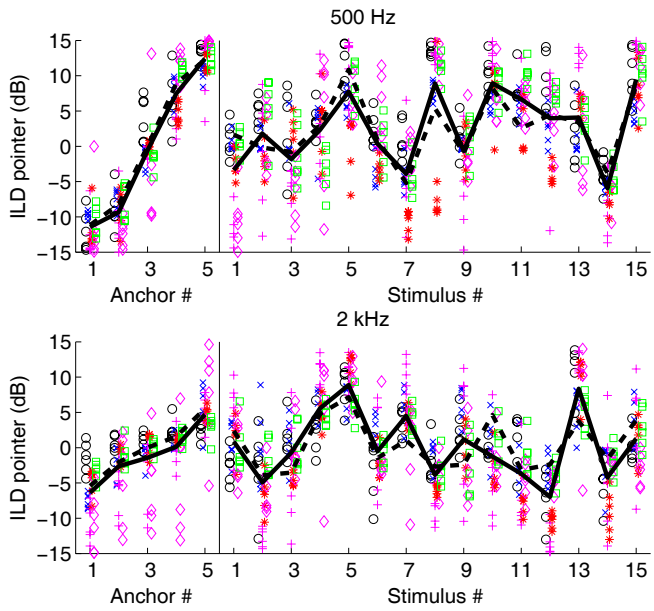

Figure 2: Results: Different symbols represent evaluations by different listeners. The continuous lines show the average results over all listeners and repetitions and the dashed lines the predictions of the cue selection model. The anchors 1–5 had ITDs of −0.5, −0.25, 0, 0.25, and 0.5 ms, respectively.

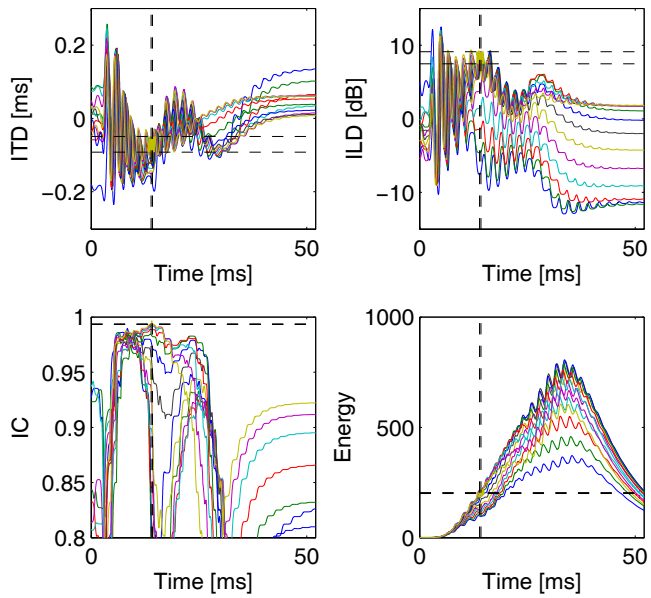

Figure 3: Selection of ITD and ILD cues: Different curves show analysis at different critical bands. The selected cues are illustrated with thick curves and dashed lines drawn over the plots. The peak IC is achieved at and the cues are selected from the critical band with a center frequency of 575 Hz.

possible off-frequency listening into account.

The ILDs were computed directly from the output of the peripheral model and the ITDs and ICs were derived from normalized cross-correlation functions. The IC was always the peak value of the cross-correlation function. For the stimuli with a center frequency of 500 Hz, best predictions were achieved when the ITD was determined such that the cross-correlation function was first weighted with a Gaussian function to emphasize central ITDs [3] and the ITD was computed as the centroid of the resulting weighted cross-correlation function. Unfortunately, the same procedure did not work well for the 2 kHz stimuli. This may be partly due to the peripheral model, which did not remove all effects of the fine structure of the signals. The cross-correlation functions exhibited periodic peaks at 0, ±0.5, ±1 ms etc. and, consequently, the centroid was always very close to 0. However, using the lag value corresponding to the peak of the unweighted cross-correlation function enabled utilizing the ITDs also at 2 kHz (see also [1]) although with somewhat reduced prediction performance compared to 500 Hz.

The cue selection was implemented following [1] such that the IC threshold was set as low as possible while still obtaining cues suggesting an unambiguous lateralization. In the current study, it was required that the standard deviation of the selected ILD cues was below the experimentally set values of 0.5 dB at 500 Hz and 2 dB at 2 kHz frequency band. The selection procedure is illustrated in Figure 3.

The mean judged pointer ILD for each stimulus was modeled according to

$$\mathrm{ILD}_{\mathrm{pointer}} = a \cdot \mathrm{ITD}_{\mathrm{sel}} + b \cdot \mathrm{ILD}_{\mathrm{sel}}\,, \qquad (1)$$

where the coefficients a and b were individually determined for the stimuli and anchors at both center frequencies using a least-squares fit over the selected cues. The results are shown in Figure 2 and they agree qualitatively with the experimental results. The optimal coefficients showed that the lateral displacement produced by the selected ITD cues of the incoherent stimuli is approximately 45% of that produced by corresponding ITDs applied to the coherent anchors. Furthermore, effectiveness of the ILDs was slightly reduced by the incoherence.

Discussion

The results show that narrowband incoherent noise bursts are lateralizable. In the modeling, highest instantaneous ICs are typically obtained during signal onsets. Hence, it is unclear whether high ICs *per se* or signal onsets trigger the auditory localization. Nevertheless, the lateralization can be fairly well predicted with the cue selection model. Further studies would be needed to separately investigate onset and IC effects.

References

[1] Faller, C. and Merimaa, J.: Source localization in complex listening situations: Selection of binaural cues based on interaural coherence. J. Acoust. Soc. Am. 116 (2004), 3075–3089.

[2] van de Par, S. and Kohlrausch, A.: A new approach to comparing binaural masking level differences at low and high frequencies. J. Acoust. Soc. Am. 101 (1997), 1671–1680.

[3] Shackleton, T. M., Meddis, R., and Hewitt, M. J.: Across frequency integration in a model of lateralization. J. Acoust. Soc. Am. 91 (1992), 2276–2279.