Acta Acustica united with Acustica
Vol. 88 (2002) 110 – 112

A Time-Domain Binaural Signal Detection Model and its Predictions for Temporal Resolution Data

J. Breebaart, S. van de Par

Philips Research Laboratories Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands

A. Kohlrausch

Philips Research Laboratories Eindhoven, Prof. Holstlaan 4, NL-5656 AA Eindhoven, The Netherlands, and Department of Human Technology Interaction, Eindhoven University of Technology, P.O. Box 513, NL-5600 MB Eindhoven, The Netherlands.

Summary

This paper discusses the application of a time-domain binaural signal-detection model in the context of estimates of the temporal resolution of the binaural auditory system. It is demonstrated that the optimal detector which is present in the model is crucial to account for specific temporal detection phenomena. In particular, the model can account for the apparent differences in the estimates of binaural time constants found with different experimental paradigms. It is argued that the differences in temporal-window estimates stem from listeners ability to listen off-time, whenever this enhances detection performance.

PACS no. 43.66.Pn, 43.66.Ba

1. Introduction

Several studies have revealed that the binaural auditory system is sluggish in its processing of interaural differences. For example, the minimum audible angle of a sound source strongly depends on its velocity [1]. Several authors have tried to describe the temporal resolution of the binaural auditory system by a single temporal averaging window. For example, experiments using time-varying interaural intensity differences (IIDs) revealed that IID detection shows a lowpass behavior with a cutoff frequency of about 20 Hz [2]. The detection of dynamic interaural time differences (ITDs) seems to be even worse [3]; dynamically-varying ITD detection has a lowpass response with a cutoff frequency of 2 to 5 Hz. Detection experiments performed using a masker with a time-varying interaural correlation show that the binaural auditory system can be described as having a time constant between 44 and 243 ms [4, 5, 6, 7], which is rather high compared to the 4 to 44 ms for monaural processing [5, 8]. The large range in the temporal estimates is in part due to the variability of the temporal resolution across subjects. However, different estimates are also obtained for one specific listener when different experimental procedures are used to determine the temporal resolution. For example, Holube *et al.* [9] estimated the time constants related to the binaural auditory system using two experimental procedures. The first method comprised the presentation of a short $S\pi$ signal in a masker with a stepwise change in interaural correlation from $+1$ to -1. If the signal was presented in the masker portion with an interaural correlation of $+1$, the stimulus condition was effectively $NoS\pi$ and hence a BMLD was observed. For a presentation of the signal in the masker portion with a correlation of -1, no BMLD was observed. The gradual change in threshold for a signal presented close to the masker-phase transition enabled the authors to estimate the binaural time constant, resulting in an temporal-window estimate with an equivalent rectangular duration (ERD) between 40 and 70 ms. The second method incorporated the presentation of an $S\pi$ signal in a masker with a sinusoidally-varying correlation. The temporal center of the signal occurred at a position corresponding to a masker correlation of $+1$. By varying the correlation-modulation frequency, Holube *et al.* estimated the ERD to be twice the value found in the previous experiment, i.e., between 90 and 120 ms. A possible explanation for these differences in the estimate of the ERD was given by Holube *et al.* They stated that the reason for this mismatch seems to be the different detection strategies employed for the various tasks that are affected by the consistency of binaural information across frequency and time. In their fitting procedure, Holube *et al.* and also Kollmeier and Gilkey [5] obtained the predicted thresholds by computing the weighted integration of the instantaneous interaural cross-correlation at the temporal center of the signal. For the sinusoidal changes in the correlation, it is likely that this detection strategy results in the highest signal-to-masker ratio, given the fact that both the temporal window and the correlation modulation are symmetric around the signal center. It is not obvious, however, that this strategy is also optimal for stepwise correlation changes; pos-

Received 25 September 2001,
accepted 27 November 2001.

sibly, off-time listening could enhance the performance. In this paper, it will be shown quantitatively that different detection strategies in the two exprimental conditions can indeed account for the discrepancy between the ERD estimates. In particular, simulations with a binaural signal detection model [10, 11, 12] demonstrated that assuming listener's possibility of off-time listening can account for the observed differences in ERD estimates.

2. Off-time listening

The model by Breebaart *et al.* was used to investigate the effect of off-time listening. This model is very suitable for this task because it incorporates an optimal detector which automatically selects the optimal time interval of a presented stimulus to derive its predictions. Since the complete model is described elsewhere, the reader is referred to the references above for a detailed model description,while a short description will be given below. Comparison with the data given by Holube *et al.* [9] is very attractive because in their fitting procedure, Holube *et al.* used a double-exponential temporal window which is exactly the same window shape as present in the binaural model.

The model consists of three stages. The first stage comprises a peripheral preprocessor, including a gammatone filterbank to simulate the spectral resolution of the basilar membrane, a simple inner haircell model and adaptation loops. The second stage consists of a binaural processor, in which the signals from the left and right ears are compared by so-called EI (Excitation-inhibition) elements, a framework which is closely related to Durlach's EC-theory [13]. The output of the EI elements is temporally smoothed by convolving the output with a double-sided exponential window with a certain equivalent rectangular duration (ERD). This smoothing process is incorporated to account for a limited binaural temporal resolution (binaural sluggishness). In the original model paper, the ERD of this window was set to 60 ms, which resulted in model simulations that closely mimic the most sensitive subjects in temporal resolution experiments [12]. Because in the present paper, we try to explain differences in estimates of binaural time constants that are derived for different experimental procedures, we have chosen to simulate individual data for one subject from the Holube *et al.* study. The subject which has a maximum BMLD that corresponds to the maximum BMLD of the model (subject CM) had an ERD estimate for sinusoidal correlation changes of 120 ms. Therefore, simulations with the same ERD of 120 ms were performed. The rationale for also applying a lower ERD value of 40 ms is explained below.

The third stage of the model consists of a central processor. The EI-type element outputs are corrupted by an additive internal noise. Subsequently, the internal representations of the external stimuli (for example in a 3 interval, forced-choice procedure) are compared to a template that consists of the average masker-alone representation from previous trials. The differences between the actual stimulus and the template are weighted and integrated both in the time and the frequency domain according to an optimal criterion. This enables the optimal detector to reduce the influence of the internal noise, and to accumulate information about the signal by adapting its observation interval (matched temporal integrator).

As stated in [12], the optimal detection moment for a stepwise correlation change may be up to 10 ms further away from the masker-phase transition than the temporal center of the signal. This means that if the listener adjusts his or her detection strategy on binaural cues present just *before or after* (depending on the stimulus configuration) the signal rather than exactly *at* the signal, better detection performance is obtained. This strategy is termed 'off-time listening'. To demonstrate the effect of off-time listening, the thresholds of subject CM in the study by Holube *et al.* will be analyzed. This subject was chosen because the maximum BMLD of this subject matches the maximum BMLD of the model. In their Figure 7, Holube *et al.* estimated the ERD for sinusoidally-varying interaural correlation at 120 ms for this particular subject, while stepwise correlation changes resulted in an ERD estimate of 40 ms. The data of this subject are shown by the triangles in the upper and lower panel of Figure 1, for the stepwise and sinusoidally-varying correlation, respectively.

The curves in Figure 1 correspond to three different conditions of model simulations.

- On-time detection only with a double-exponential window with an ERD of 40 ms. This simulation was performed to check whether the data with stepwise correlation changes of Holube *et al.* could be replicated by the model. Results are shown by the dotted curves.
- Optimal detection (i.e., with the possibility of off-time listening) following temporal averaging with a double-exponential window with an ERD of 40 ms. This simulation gives insight whether off-time listening indeed results in increased detection performance. Results are shown by the solid curves.
- Optimal detection following temporal averaging with a window ERD of 120 ms. This simulation tends to provide evidence that the decrease in detection performance by the degraded temporal resolution can be compensated by off-time listening. The results are shown by the dashed curves.

If the stepwise correlation changes are considered (upper panel of Figure 1), the on-time simulation with a 40-ms-ERD window gives a good fit to the data, demonstrating that we can indeed replicate the original analysis by Holube *et al.* (dotted curve). If off-time listening is enabled and the same temporal window is used (solid curve), the simulated thresholds decrease by several dB indicating that indeed off-time listening can increase the detection performance. The dashed curve shows that this advantage can be compensated by enlarging the ERD of the temporal-averaging window of the model. Hence from these results alone, it is difficult to discriminate between a system which has a window ERD of 120 ms combined with the ability to listen off-time, and a system with a window ERD of 40 ms which can only listen on-time.

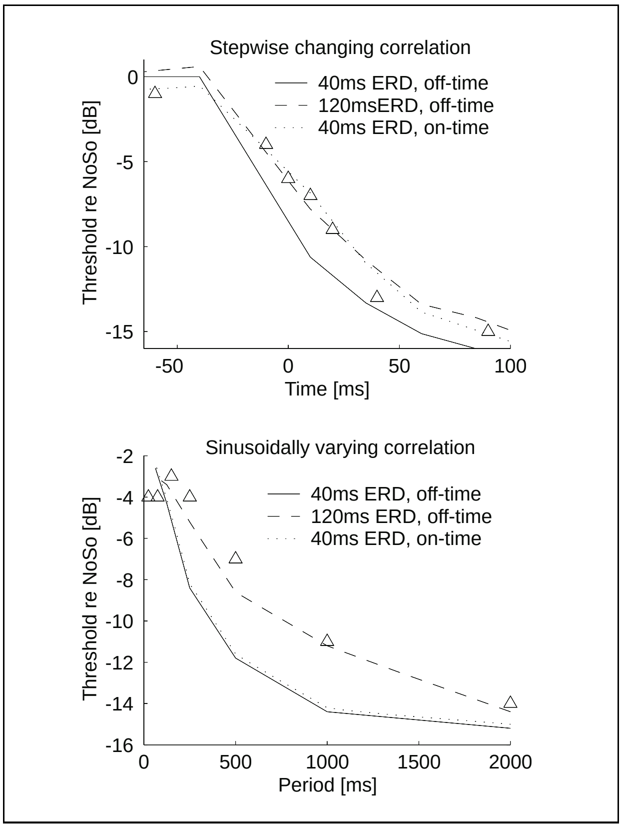

Figure 1. Experimental data adapted from Holube *et al.* [9] for stepwise masker correlation changes (upper panel) and sinusoidally-varying interaural masker correlation (lower panel). Thresholds re NoSo are shown as a function of the temporal center of the signal relative to the masker phase-transition (upper panel) and as a function of the modulation period (lower panel). The curves represent simulations with different model settings (see text).

The lower panel of Figure 1 shows thresholds re NoSo for a sinusoidally-varying masker correlation. In this condition, off-time listening does not improve the detection performance and optimal detection can be obtained *on-time*Thus off-time and on-time predictions are the same (solid and dotted curve). Furthermore, the simulation with a window ERD of 40 ms (solid curve) clearly overestimates the detection performance of the subjects. On the other hand, the simulation with a window ERD of 120 ms fits the data quite well. Thus, both experiments can be accurately described by a model which includes a fixed temporal window with an ERD of 120 ms which is able to listen off-time if such a process enhances the detection performance. Furthermore, it seems that the fitting procedure used by Holube *et al.* *underestimates* the time constants of the binaural auditory system by excluding the possibility of off-time listening in experiments with a stepwise correlation change. Finally, it should be mentioned that although the model predictions were obtained for masker correlations changing from $+1$ to -1, the model can also account for similar effects obtained with correlation changes from -1 to $+1$ (see [12] for the corresponding model simulations).

3. Conclusions

The temporal resolution data obtained by Holube *et al.* suggested that different experimental conditions result in different estimates of the binaural auditory time constant. It was argued that these differences result from different detection strategies. More precisely, the data suggest that listeners are able to listen ’off-time’ if such a process enhances the detection performance. Therefore, experiments that aim at estimating time constants should either prevent listeners to listen off-time or include off-time listening in the analysis of the data.

Acknowledgement

The authors wish to thank the two anonymous reviewers, and Ronald Aarts and Nicolle van Schijndel for their valuable comments in improving the original manuscript.

References

[1] D. R. Perrott, A. D. Musicant: Minimum auditory movement angle: binaural localization of moving sound sources. J. Acoust. Soc. Am. **62** (1977) 1463–1466.

[2] D. W. Grantham: Discrimination of dynamic interaural intensity differences. J. Acoust. Soc. Am. **76** (1984) 71–76.

[3] D. W. Grantham, F. L. Wightman: Detectability of varying interaural temporal differences. J. Acoust. Soc. Am. **63** (1978) 511–523.

[4] D. W. Grantham, F. L. Wightman: Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J. Acoust. Soc. Am. **65** (1979) 1509–1517.

[5] B. Kollmeier, R. H. Gilkey: Binaural forward and backward masking: evidence for sluggishness in binaural detection. J. Acoust. Soc. Am. **87** (1990) 1709–1719.

[6] J. F. Culling, Q. Summerfield: Measurements of the binaural temporal window using a detection task. J. Acoust. Soc. Am. **103** (1998) 3540–3553.

[7] M. A. Akeroyd, A. Q. Summerfield: A binaural analog of gap detection. J. Acoust. Soc. Am. **105** (1999) 2807–2820.

[8] C. J. Plack, B. C. J. Moore: Temporal window shape as a function of frequency and level. J. Acoust. Soc. Am. **87** (1990) 2178–2187.

[9] I. Holube, M. Kinkel, B. Kollmeier: Binaural and monaural auditory filter bandwidths and time constants in probe tone detection experiments. J. Acoust. Soc. Am. **104** (1998) 2412–2425.

[10] J. Breebaart, S. van de Par, A. Kohlrausch: Binaural processing model based on contralateral inhibition. I. Model setup. J. Acoust. Soc. Am. **110** (2001) 1074–1088.

[11] J. Breebaart, S. van de Par, A. Kohlrausch: Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. J. Acoust. Soc. Am. **110** (2001) 1089–1104.

[12] J. Breebaart, S. van de Par, A. Kohlrausch: Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters. J. Acoust. Soc. Am. **110** (2001) 1105–1117.

[13] N. I. Durlach: Equalization and cancellation theory of binaural masking-level differences. J. Acoust. Soc. Am. **35** (1963) 1206–1218.