



# Tone detection thresholds in interaurally delayed noise of different bandwidths

Mathias Dietz\*, Jörg Encke, Kristin I Bracklo, and Stephan D Ewert

Department für Medizinische Physik und Akustik and Cluster of Excellence “Hearing4all”, Universität Oldenburg, 26111 Oldenburg, Germany

Received 18 May 2021, Accepted 1 December 2021

**Abstract** – Differences between the interaural phase of a noise and a target tone improve detection thresholds. The maximum masking release is obtained for detecting an antiphasic tone ( $S_\pi$ ) in diotic noise ( $N_0$ ). It has been shown in several studies that this benefit gradually declines as an interaural time delay (ITD) is applied to the noise. This decline has been attributed to the reduced interaural coherence of the noise. Here, we report detection thresholds for a 500 Hz tone in masking noise with ITDs up to 8 ms and bandwidths from 25 to 1000 Hz. Reducing the noise bandwidth from 100 to 50 and 25 Hz increased the masking release for 8-ms ITD, as expected for increasing temporal coherence with decreasing bandwidth. For bandwidths of 100–1000 Hz no significant difference in masking release was observed. Detection thresholds with these wider-band noises had an ITD dependence that is fully described by the temporal coherence imposed by the typical monaurally determined auditory-filter bandwidth. A binaural model based on interaural phase-difference fluctuations accounts for the data without using delay lines.

**Keywords:** Binaural hearing, Binaural unmasking, Interaural phase difference, Interaural coherence

## 1 Introduction

The human binaural system can exploit differences between the interaural phase of a masker noise and a target tone to improve detection thresholds [1]. The maximum binaural masking level difference (BMLD) is obtained for detecting an antiphasic tone ( $S_\pi$ ) in in-phase noise ( $N_0$ ). If an interaural time difference (ITD) is applied to the noise the  $S_\pi$  detection thresholds increase. Two aspects have been identified to contribute [2]. First, the noise ITD directly translates into an interaural phase difference (IPD). If the ITD is exactly half of the tone period, it is as antiphasic as the tone around the tone frequency, and the binaural benefit is diminished [2]. This can be avoided by applying the ITD to both tone and noise [3] or by looking at ITDs that are an integer multiple of the tone period, as discussed in [2]. The second aspect is that the binaural benefit gradually declines with increasing ITD even if the first periodic issue is avoided. Two different mechanisms have been proposed to account for the latter gradual decline, which is in the focus of the present study.

One mechanism, proposed by Langford and Jeffress [2], attributed the reduction of the BMLD with increasing noise ITD to a reduction in what they referred to as “interaural correlation of the noise”. A more contemporary wording

for their quantity is “normalized cross-correlation coefficient”, i.e. the value of the normalized cross-correlation function at  $\tau = 0$ .

The other mechanism, also proposed by Jeffress [4], but previously to the above, is that using internal-time-delay elements (“delay lines”) the auditory system has access to more of the cross-correlation function than just the cross-correlation coefficient at  $\tau = 0$ . Such circuitry has indeed been found in the barn owl, where left- and right-sided inputs propagate along counterdirected axons [5]. Coincidence-detecting neurons along the axonal delays effectively cross correlate the inputs at different values of  $\tau$ . An ideal delay line could perfectly compensate any external noise ITD by an opposed internal delay ( $\tau = -\text{ITD}$ ), allowing for maximum BMLDs even at large noise ITDs [6]. It is, however, reasonable to assume that such a compensation mechanism introduces errors for increasing internal delays. This increase in error is commonly simulated as a decrease of the density of correlating elements with increasing internal delay. This relationship is captured by the  $p(\tau)$  function [7, 8]. Based on this second mechanism, models use  $p(\tau)$  as a fitting parameter, i.e. they estimate the delay-line length and potency from the decline of the BMLD with noise ITD [7–10].

For the first “cross-correlation coefficient”-based concept, however, this degree of freedom does not exist. In this case, the cross correlation, or more generally speaking the complex-valued temporal coherence of the analytical signal  $\gamma$  [11],

\*Corresponding author: [m.dietz@uol.de](mailto:m.dietz@uol.de)

$$\gamma(\tau) = \frac{\langle n(t+\tau)n^*(t) \rangle}{\langle |n(t)|^2 \rangle}. \quad (1)$$

is solely determined by the spectrum of the noise [2, 6] and is proportional to the inverse Fourier transform  $F^{-1}$  of its power spectral density  $n'$  (Wiener–Khinchin theorem):

$$\gamma(\tau) \propto F^{-1}(n'(f)). \quad (2)$$

This proportionality means that the bandwidth of the input signal determines the decay of  $|\gamma(\tau)|$ : the broader the spectrum, the shorter the temporal coherence. The maximum bandwidth observed by the binaural system is presumed to be limited by some form of band-pass filter in the periphery. Consequently, the filter-bandwidth effective at the input to the binaural interaction ultimately determines how binaural unmasking depends on the noise ITD. If the stimulus bandwidth is much wider than the filter bandwidth, only the filter bandwidth determines the ITD dependence. For noise bandwidths similar to the filter bandwidth, the filter shape plays a certain role. For narrowband stimuli, the decline would be even more gradual and caused by the stimulus bandwidth. As a first hypothesis, it seems reasonable to assume that the effective bandwidth at the binaural stage matches those from monaural estimates (e.g., 79 Hz ERB at 500 Hz [12]). Previous studies indeed already tried to determine the effective bandwidth based on binaural unmasking data. Rabiner et al. [6] found that their experimental data could be best accounted for using an 85-Hz wide (at  $-3$  dB) triangular filter. Langford and Jeffress [2] coarsely estimated a 100-Hz bandwidth but without specifying the filter shape and bandwidth definition. Both estimates are larger than but close to the monaural estimates.

The goal of this study is to revisit if and to what extent the decline in binaural unmasking as a function of noise ITD can be explained solely based on the decline of temporal coherence in a simple, “minimalistic” binaural model. The model uses fluctuations of the interaural phase difference [13, 14] resembling a physiologically plausible feature that might be extracted from a neural representation of the signal [15]. This IPD fluctuation metric is directly related to the correlation coefficient, i.e., to the degree of coherence [16, 17]. Early attempts [2, 6] connecting noise ITD and temporal coherence appear promising, but have not been followed up, to test whether filter bandwidths and filter shapes that are more commonly used in recent models can quantitatively account for the data. If the simulated decline is faster than the experimental decay, delay lines could offer an explanation, compensating for the external ITD and thus increasing the coherence at the level of binaural interaction.

To complicate the argument, there has been controversy related to the effective processing bandwidth of the binaural system, with some studies suggesting it is larger than the monaural filter bandwidth, at least for certain complex maskers [8, 18, 19]. However, there is growing consensus that, at least for the simplistic band-pass filtered noise investigated in the present study, the binaural filter bandwidth is not wider [9, 10, 20–24].

## 2 Experiment

### 2.1 Participants

Ten young normal-hearing volunteers (21–33 years, median 24 years, 5 male, 5 female) were recruited – all university students. Most subjects had some experience in speech-in-noise tests. To our knowledge, no subject had prior experience in dichotic tone-in-noise detection. All subjects received at least 90 min of training prior to data collection. All audiometric thresholds were equal to or less than 15 dB HL from 125 to 10 000 Hz, averages across the test frequencies were less than or equal to 5 dB HL, and differences across ears did not exceed 5 dB in both pure-tone average and at 500 Hz. One reason for these relatively strict inclusion criteria is that it has been shown recently that subjects with a slight (sub-clinical) hearing loss have a reduced binaural release from masking [25]. At the test frequency of 500 Hz, no subject had a threshold  $> 5$  dB HL in either ear. Subjects were compensated for their time.

### 2.2 Apparatus

Stimuli were generated digitally at a sampling rate of 48 kHz in MATLAB (MathWorks, Natick, MA, United States) using the AFC software package [26] for MATLAB and presented via an RME (Haimhausen, Germany) Fireface UC USB sound card, over a Sennheiser (Wedemark, Germany) HDA 650 circumaural headphone, calibrated at 500 Hz. The subjects were seated in a double-walled sound booth.

### 2.3 Stimuli

Gaussian noise was band-pass filtered by cutting out spectral components outside the pass-band. All noises were arithmetically centered at 500 Hz and the bandwidths were 25, 50, 100, 150, 200 or 1000 Hz. Noises had a duration of 380 ms including 20-ms raised cosine onset and offset ramps. The noise level was kept at a constant spectrum level of 45.5 dB relative to 20  $\mu$ Pa. Fully correlated noises were presented with ITDs of 0, 2, 4, or 8 ms, or were interaurally uncorrelated. The ITD was applied prior to gating. ITDs were chosen in multiples of the cycle duration at the 500-Hz center frequency, ensuring zero interaural phase difference (IPD) at 500 Hz.

Target tones had a frequency of 500 Hz and a duration of 300 ms, again including 20-ms raised cosine on- and offset ramps. Tones were always presented temporally centered in the noise. Target tones were either interaurally in phase ( $S_0$ ) or antiphasic ( $S_\pi$ ).

### 2.4 Procedure

A 3-interval, 3-alternative forced-choice procedure was employed, with two noise-only reference intervals ( $N_r$ ) and one target interval including both signal and noise. Subjects selected an interval by pressing the respective number key on a computer keyboard. Feedback was provided.

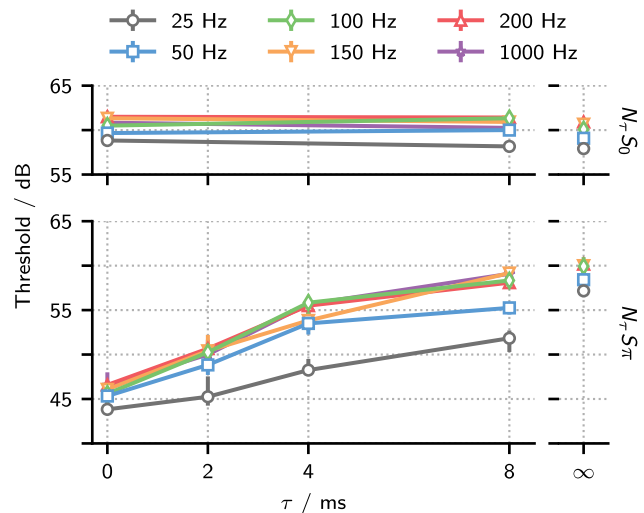
The signal level, initially 65 dB SPL, was adaptively changed in a 2-down, 1-up staircase procedure, aiming at the 70.7% correct rate [27]. The step size of 4 dB was reduced to 2 and 1 dB after the second and fourth reversal, respectively. After a total of 10 reversals each run was terminated, and the average was taken across the last 6 reversals.

$S_0$  and  $S_\pi$  conditions were separated into two independent experiments presented in blocks but without a specified order. For  $S_0$  conditions only noise ITDs of 0 and 8 ms, as well as uncorrelated noise, were measured. For both the  $S_0$  and  $S_\pi$  conditions the same measurement order principles were applied: For each randomly chosen bandwidth block all noise ITDs were measured once in random order. To allow for an “acclimatization” to the new bandwidth, two training runs were included at the beginning of each bandwidth block. After all conditions, i.e., all six bandwidth blocks, were tested, the procedure was repeated with new random orders until all conditions were tested 4 times. The study was approved by the Ethics committee of the University of Oldenburg.

## 2.5 Results and discussion

Three of the ten listeners were not able to obtain  $N_0S_\pi$  thresholds below +8 dB above the masker spectrum level in the 100-Hz bandwidth noise, while the seven other subjects had thresholds less than or equal to +2 dB in their measurements. Comparable thresholds from other studies are in the range of -3 dB [20] to +3 dB [3]. These three listeners had excellent audiograms but performed quite poorly and in fact worse than the group average of less-sensitive listeners in [25]. Two of the less-sensitive listeners were further tested in a sensitivity-optimized threshold ITD task, as in [28]. Their  $>100 \mu\text{s}$  thresholds were larger than any of 52 un-trained listeners [28], and thus they were considered to be outliers. To be able to compare the data to previous studies and to meaningfully apply statistics based on normally distributed values, only data from the other seven subjects are reported.

The experimental data are shown in Figure 1 as a function of the noise ITD from 0 to 8 ms and for interaurally uncorrelated ( $N_u$ ) noise ( $\infty$ ). Symbols depict the median detection thresholds across the seven subjects and error bars the interquartile range. Different symbols are used for different noise bandwidths. The lower panel shows thresholds for the  $S_\pi$  condition.  $N_0S_\pi$  thresholds (left-hand data points) were similar to the masker spectrum level and for large bandwidths were virtually identical to a large and consistent body of literature [20, 29]. As expected,  $S_\pi$  detection thresholds increase with increasing noise ITD. The thresholds obtained with 100- to 1000-Hz wide noise are virtually identical with each other and to those obtained by Langford and Jeffress [2]. Smaller bandwidths of 50 and 25 Hz resulted in slightly lower  $N_uS_\pi$  and  $N_0S_\pi$  thresholds and the latter increased more gradually with increasing noise ITD. Qualitatively, all of these observations were expected as a direct consequence of the increasing temporal coherence in the acoustic stimulus for

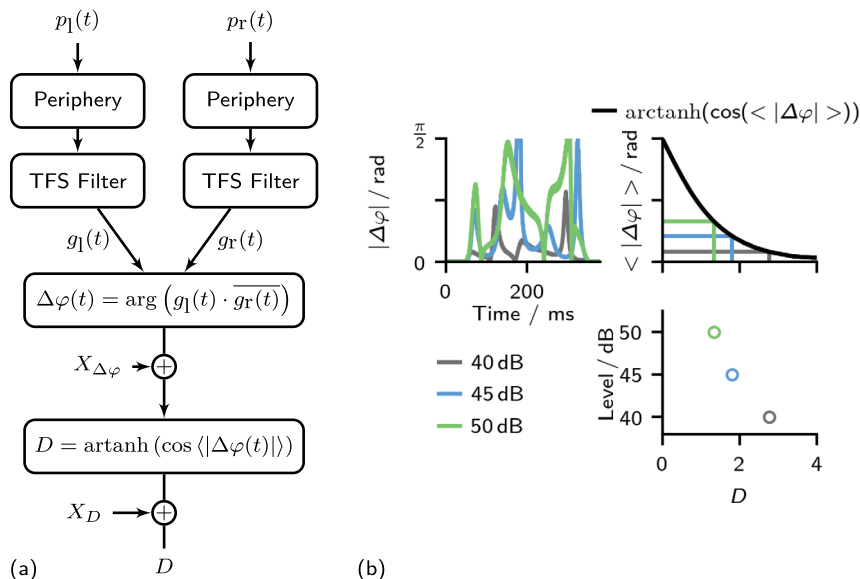


**Figure 1.**  $N_\tau S_0$  detection thresholds (upper panel) and dichotic  $N_\tau S_\pi$  thresholds (lower panel) as a function of noise ITD in ms. The separated data points at the right-hand side are for uncorrelated noise ( $N_u$ ). Different symbols (color online) are for the different noise bandwidths ranging from 25 Hz to 1000 Hz.

decreasing bandwidth. For narrow bandwidths, up to about 100 Hz, previously reported thresholds for  $N_0S_\pi$  are less consistent across studies. The  $N_0S_\pi$  thresholds of the present study are lower than those reported by Bernstein and Trahiotis [3], similar to those by van der Heijden and Trahiotis [8], but overall higher than in the majority of studies focused on thresholds of highly trained listeners obtained with diotic maskers [20, 30]. We speculate that mixing in experimental conditions with nonzero ITDs makes it harder for the subjects to fully train on the particularly subtle cues with narrow-band, i.e. tonal maskers and zero ITD.

$S_0$  thresholds are shown in the upper panel of Figure 1. These data were not the focus of the current study and were recorded only for ITDs of 0 and 8 ms as well as for the uncorrelated noise to estimate the observed BMLD. The BMLD was on average 14.8 dB, without noise ITD, independent of bandwidth, and was reduced to about 6.3 dB and 4.7 dB for 25-Hz and 50-Hz bandwidth at 8-ms noise ITD, respectively. For larger bandwidths the BMLD at 8-ms noise ITD was only about 2 dB, in line with [2], and for uncorrelated noise it was on average 1 dB.

To further assess the effect of bandwidth on the  $N_\tau S_\pi$  thresholds, a two-way repeated-measures ANOVA [noise ITD (5)  $\times$  bandwidth (6)] was performed, showing a significant main effect of noise ITD [ $F(4, 24) = 464.30$ ,  $p < 0.001$ ], bandwidth [ $F(5, 30) = 21.53$ ,  $p < 0.001$ ], and a significant interaction [ $F(20, 120) = 6.47$ ,  $p < 0.001$ ]. Post-hoc pair-wise comparisons (Bonferroni corrected) for the marginal means showed that thresholds at all noise ITDs were significantly different ( $p < 0.01$ ) from each other. For the bandwidths, only the data for 25 Hz were significantly different from all other bandwidths ( $p < 0.05$ ), and the data for 50 Hz were significantly different from



**Figure 2.** (a) Schematic of the proposed model. (b) Example of the decision stage for  $N_0 S_\pi$  stimuli at different signal levels. The top left shows the absolute value of the instantaneous IPD over time. Increasing the target level (with an IPD of  $\pi$ ) increases fluctuations and thus the mean absolute IPD ( $y$ -axis top right). Taking the cosine of the mean IPD results in smaller output values with increasing IPD fluctuations. The decision variable  $D$  results directly from the cosine, subject to a Fisher’s-Z transform, so that lower values indicate a stronger signal prevalence (bottom right).

200 Hz ( $p < 0.05$ ). At the largest ITD of 8 ms, the data for 25 Hz were different ( $p < 0.05$ ) from all other bandwidths, except for 50 Hz, and the data for 50 Hz were significantly different ( $p < 0.05$ ) from all other bandwidths except for 25 and 200 Hz. Such deviations in the pattern of differences from the marginal mean and between the different ITDs are the source of the significant interaction term. Taken together, the 25-Hz data, and the 50-Hz data to a lesser degree, were different from the data for the other bandwidths, while the data for 100-Hz bandwidth and above showed no significant differences.

### 3 Model predictions

#### 3.1 Model description

The front end of the model employed here was essentially identical to the IPD model [13, 14] and illustrated in Figure 2a. Peripheral processing was done with very simplistic signal processing stages typical for this class of models. Parameters for the peripheral stages had been derived prior to [13, 14] primarily by fitting psychoacoustic data. Band-pass filtering was simulated with a 4th-order Gammatone filter centered at the target frequency of 500 Hz, with equivalent rectangular bandwidths of 50, 60, 79, 100, and 130 Hz in the implementation of [31]. A static compression was included by taking the signal to a power of 0.4. Loss of fine-structure ITD sensitivity was coarsely modelled by half-wave rectification and subsequent low-pass filtering with a 5th-order Butterworth filter with a 770-Hz cutoff frequency).

For IPD extraction, the phases of left and right signals must be known by definition. The phase is extracted from

the peripheral representation by applying a second, broader Gammatone filter (2nd-order, 167-Hz bandwidth), referred to as temporal fine-structure (TFS) filter, again centered at 500 Hz. From the complex-valued output of this TFS filter,  $g(t)$ , the argument is the phase. The TFS filter effectively reverts some effects of the half-wave rectification, similar to the band-pass characteristic of ITD-sensitive neurons in the medial superior olive [32]. In principle, for the purpose of the present study, the phase could have been obtained directly from the first Gammatone filter, however, the peripheral stage and the TFS filter were kept as in the IPD model [13, 14] to stay in the conceptual framework of auditory pathway models.

The instantaneous IPD,  $\Delta\varphi(t)$ , was then derived by subtracting the phases from the left and the right signal, or, equivalently, by first multiplying the left signal and the complex conjugate of the right signal and then taking the argument from the product. A phase jitter,  $X_{\Delta\varphi}$ , in the form of Gaussian noise was added to the IPD as a limiting factor of binaural sensitivity, qualitatively corresponding to the time-equalization jitter introduced by Durlach [33] or a combined monaural and binaural time jitter [34].

Adding an  $S_\pi$  tone to more intense diotic noise causes the instantaneous IPD to fluctuate around zero. This fluctuation has previously been suggested as the detection cue [13, 17, 35–37]. In more general terms, it is assumed in the model that the target can be detected if the average fluctuation of the IPD in the target interval can be discriminated from the average fluctuation of the IPD in the noise-alone intervals.

In the current study, the long term average IPD of both target and reference intervals are always zero:  $\langle \Delta\varphi \rangle = 0$ . Without loss of generality, the present study employs a

simplified implementation of the model and only detects deviations from zero IPD, i.e. temporal averaging of the modulus across the entire observation interval:  $\langle |\Delta\phi| \rangle$ . This value ranged from 0, in case of a diotic stimulus and no internal noise, to  $\pi$ , for an interaurally antiphasic stimulus. For interaurally uncorrelated noise an average value of  $\pi/2$  was obtained, resulting from a uniform distribution of  $|\Delta\phi|$  in the range 0 to  $\pi$  in that case. A cosine mapping, projected the values to the interval 1 for no IPD, to  $-1$  for an antiphasic stimulus, and to 0 for an interaurally uncorrelated noise. In this simplified version, which could only compare intervals with no offset IPD, the internal variable  $\cos\langle |\Delta\phi| \rangle$  is practically identical to the interaural correlation coefficient, commonly used for similar purposes [10, 29]. This term was then Fisher-Z transformed (Fig. 2b), again identical to the comprehensive binaural detection model by Bernstein and Trahiotis [29]. Last, a detector noise  $X_D$ , was added, yielding the decision variable  $D$ :

$$D = \operatorname{arctanh}(\cos\langle |\Delta\phi| \rangle) + x_D. \quad (3)$$

The  $\operatorname{arctanh}$  Fisher-Z transformation expands differences near 1 and  $-1$  and has previously been employed for correlation-based decision metrics [38] as well as for dichotic tone in noise detection [28]. In combination with the detector noise  $X_D$ , an increased sensitivity to interaural correlation differences near 1 and  $-1$  was obtained in combination with decreased sensitivity close to 0 (uncorrelated) as observed in, e.g. [39]. The model back-end was an artificial observer that evaluated the same three intervals in the same adaptive procedure as the listeners and selected the interval with the smallest decision variable  $D$  [9, 40].

For each of the five filter bandwidths the two internal noise parameters were fit to perfectly match the two 1000-Hz noise bandwidth thresholds at 0 and 8-ms ITD. First  $X_D$  was varied to match the model to the average experimental 8-ms threshold, then the IPD noise was varied to fit the 0-ms threshold. Due to the codependence, the second step caused a small change in the 8-ms thresholds, so a second optimization iteration was conducted. After this fitting, the artificial observer ran 200 runs in each condition.

### 3.2 Model simulations

Figure 3 shows the model results for the five different filter bandwidths. For comparison, the experimental data is plotted in the background as dashed lines. In the model, the thresholds obtained at 100-Hz bandwidth (diamonds) differed slightly from the 1000-Hz condition. Both absolute values and the bandwidth-dependent increase with noise ITD were quite accurately captured by the model, especially with auditory-filter bandwidths of 50–100-Hz ERB. With the 130-Hz filter, the narrow- and broad-band noise data cannot be simulated at the same time. Relatedly, the 130-Hz filter always results in too large threshold differences between the 100-Hz bandwidth conditions and the 150-, 200-, and 1000-Hz conditions. The influence of filter bandwidth on noise bandwidth shown in Figure 3 may be non-intuitive, because the broadband noise-conditions appear unaffected by the filter bandwidth and the 25-Hz

thresholds are most affected. This influence is caused by the different internal noise parameters (see Tab. 1). With identical internal noise parameters the thresholds for the 25-Hz conditions are unaffected by filter bandwidth, and changes occur at the thresholds at larger noise bandwidths. Diotic conditions were not modeled, given that monaural cues are not captured by the purely binaural model. This limitation also explains the fact that the model thresholds were slightly too high for uncorrelated noise (Fig. 3 bottom, light gray): subjects apparently use a combination of weak binaural cues and weak monaural cues in these conditions.

The two internal-noise parameters influenced the model output in the following way: The internal IPD noise ( $X_{\Delta\phi}$ ) mostly determined the threshold at  $\tau = 0$  and, the slope and curvature of the threshold functions for increasing ITD. The decision noise ( $X_D$ ) determined the overall performance, effectively shifting the functions towards higher or lower thresholds.

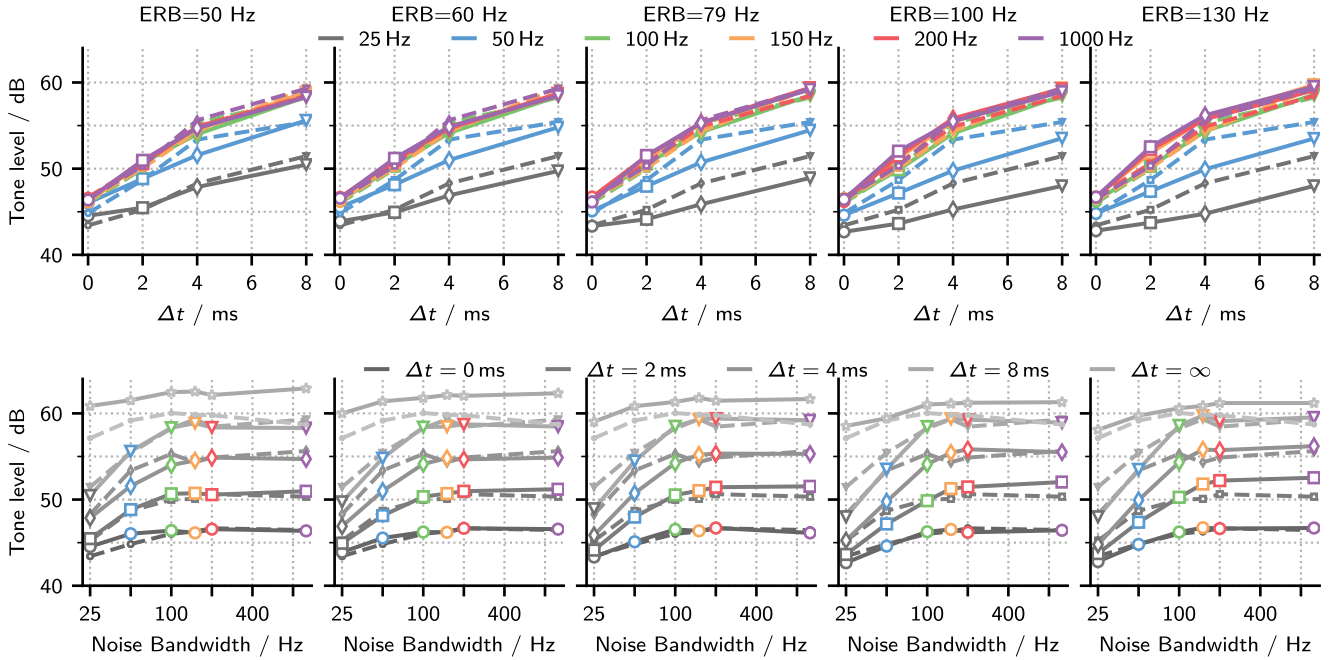
### 3.3 Application of the model to data from the literature

The suggested model was additionally tested in comparison to data from the literature for interaural-correlation discrimination [39] and binaural unmasking with arbitrary group delays [6, 41], i.e., a noise ITD that leaves the TFS ITD fixed at zero as the model implementation requires.

The current model can be directly applied to interaural correlation discrimination tasks. The internal-noise parameter value was also kept unchanged. Results are shown in Figure 4a, again with the model operating as an artificial observer. Following the approach of [39], the  $d'$ s were calculated between selected interaural correlations, as given in their Table 1.  $d'$  for the measured values of interaural correlation  $\rho$  and the fully correlated stimulus ( $d'_{(\rho,1)}$ ) were then calculated by systematically summing over the calculated  $d'$ , as described in [39]. This process resulted in multiple approximations of  $d'_{(\rho,1)}$  (e.g.  $d'_{(0,0.8)} + d'_{(0.8,1)} = d'_{(0,1)}$  and  $d'_{(0,0.5)} + d'_{(0.5,0.9)} + d'_{(0.9,1)} = d'_{(0,1)}$ ) all of which are given in Figure 4a. The model was mostly able to reproduce the data. Only for uncorrelated noise the model slightly outperforms the average experimental data.

Additionally, the model was employed to simulate detection thresholds obtained with noise maskers with an ITD in the range of 0–7.8 ms and an additional phase shift to adjust the resulting interaural phase difference at the target frequency of 500 Hz to zero, i.e., a pure group delay [6]. Results are shown in Figure 4b. While no adjustment of the model parameters was necessary to reproduce the data of [39], an increase in the standard deviation of the IPD noise to  $\sigma_{\Delta\phi} = 0.45$  rad was required to simulate this data set. This increase in internal noise might be attributed to a shorter stimulus duration (while no quantification of the stimulus duration is given in [6] the stimuli were described as “short”).

Thirdly, [41] measured tone-detection thresholds in 50-Hz and 400-Hz wide noise maskers with a fixed group delay of 1.5 ms. Relative to their respective  $N_0S_\pi$  reference, thresholds did not increase for the 50-Hz condition, but they increased by about 5 dB in their 400-Hz condition.



**Figure 3.** Model predictions (solid lines with symbols) and experimental data (dashed lines). The five columns show predictions obtained with five different filter bandwidths. In the top row, thresholds are plotted as in Figure 1. In the bottom row, the same data are plotted as a function of noise bandwidth with different lines representing the different noise ITDs. The symbol color is kept from the bandwidth color-coding in the top row. Data for uncorrelated noise maskers is only shown in the bottom panels in light gray. No error bars are shown for the simulations, because the standard error of the mean was  $<0.5$  dB for all conditions.

**Table 1.** Internal noise parameters,  $R^2$  correlation between the 24 experimental and simulated  $N_s S_\pi$  thresholds and root-mean-square error (RMSE) for the five different filter bandwidths.

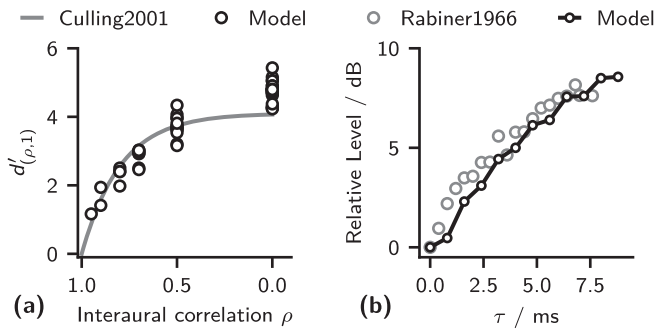
ERB	50 Hz	60 Hz	79 Hz	100 Hz	130 Hz
$X_{\Delta\phi}$	0.25 rad	0.28 rad	0.31 rad	0.34 rad	0.37 rad
$X_D$	0.58	0.475	0.385	0.31	0.28
R	0.98	0.97	0.95	0.91	0.89
RMSE	0.72 dB	0.85 dB	1.09 dB	1.48 dB	1.58 dB

Their model could not account for their observation, whereas the present model predicts a 2-dB increase for the 50-Hz condition and a 4-dB increase for the 400-Hz condition.

## 4 Discussion

The experimental data and simulations presented in this study assessed the effect of tone-detection thresholds as a combined function of noise ITD and noise bandwidth. The lower the bandwidth of the noise, and thus the higher the temporal coherence of the noise, the smaller the impact of noise ITD on the threshold. The experimental data systematically extends the classical ITD dependence for broadband noise [2] by adding five narrower bandwidths of 25–200 Hz. The experiment extended [10] both the number of bandwidths tested (6 instead of 2) and the ITD range (8 ms and infinity instead of 3 ms).

In order to interpret the experimental results, a minimalistic phenomenological model based on the IPD-model front end [13] was employed. The most relevant aspect of the model is the peripheral-filter bandwidth, which was varied from 50- to 130-Hz ERB. The standard setting with 79-Hz wide filters most closely resembles psychoacoustic estimates of an auditory-filter band centered at 500 Hz [12]. The model has two internal-noise parameters: (1) a phase jitter, and (2) a detection uncertainty. The model reproduces all trends in the data and can quantitatively account for all data except for uncorrelated noise, especially



**Figure 4.** (a) Comparison between the results of the correlation-discrimination task of Culling et al. [39] (grey line shows their fit for the average subject) and the model performance in the same task (b). Comparison between the model results (79-Hz filter bandwidth) and the data from Rabiner et al. [6], where the  $x$ -axis shows the group delay and the  $y$ -axis the associated threshold level relative to the  $\tau = 0$  condition.

with narrow peripheral filters up to the 79-Hz standard width.

Most information about the filter bandwidth appears to be provided by responses to noise ITDs of 2 and 4 ms, where the ITD causes a partial decorrelation, but thresholds are still more than 4 dB below thresholds with uncorrelated noise. This decorrelation increases the threshold differences between narrow- and broad-band maskers observed in zero-ITD noise. The model versions with ERBs of 100 and 130 Hz fail to quantitatively predict this difference by 3 and 4 dB, respectively. Even with the  $ERB = 79$  Hz standard setting the model overestimates the difference. Narrower peripheral filters would be a possible explanation and have previously been suggested based on otoacoustic emissions [42]. At 8-ms noise ITD, five of six thresholds are still more than 1 dB below the respective threshold for uncorrelated noise, hinting at a residual correlation.

The model uses the average absolute value of the instantaneous IPD as its decision metric. This metric is used with the underlying assumption that IPD fluctuations are a cue in binaural tone-in-noise detection tasks [35, 36], possibly even a more accurate cue than interaural correlation [37]. The binaural comparator element(s) can thus distinguish between, e.g., interaurally uncorrelated inputs causing large IPD fluctuations and a constant IPD of  $90^\circ$ . For a conventional cross-correlation element, both stimuli generate the same output: zero correlation. For the present dataset, all interaurally delayed noises have a long-term average IPD of zero, i.e., the cue is the IPD deviation from zero. Therefore, for the present data set, our IPD-deviation-from-zero implementation is practically equivalent to the correlation-based model of Bernstein and Trahiotis [10, 29], when looking only at the output of the center ( $\tau = 0$ ) delay-line element. Bernstein and Trahiotis [29] accounted for almost all of the data from six seminal studies without requiring delay lines, i.e., with the before-mentioned  $\tau = 0$  element. They only employed delay lines for tone detection in interaurally delayed noise. However, with the primary difference in the implementation of the internal-noise parameters, we see no reason why their model should not account for our interaurally delayed noise data without using non-zero delays.

It is beyond the scope of the present study to simulate data sets with non-zero mean IPDs, such as [2, 8], but first attempts with a manual setting of the mean IPD appear encouraging. Equivalently, models based on the correlation coefficient could account for the binaural advantage offered by such noise, without requiring a delay line, if they used the complex-valued correlation coefficient, i.e. interaural coherence [11], rather than the real-valued correlation.

The near-equivalence of the present IPD model and models based on interaural correlation or coherence has further been illustrated by simulating interaural-correlation discrimination data [39] without changing model parameters. For the present model it does not make a difference if the masker-inherent IPD fluctuation, or synonymously the reduction in masker coherence, is introduced by an ITD or by adding uncorrelated noise, which is in line with [2, 6]. Moreover, it was shown that the current model is able to account quantitatively for  $S_\pi$  detection thresholds in

noises with arbitrary group delays when the IPD (i.e. the TFS ITD) at 500 Hz was fixed at zero. For noises much wider than the filter bandwidth, the model predicted a threshold increase of just under 3 dB per millisecond ITD, more than observed in [6] and less than observed in [41]. For 50-Hz wide noise, the increase was closer to 2 dB/ms. Many delay-line-based approaches fail to account for this bandwidth dependence because sensitivity is dictated by the bandwidth-independent delay-line potency  $\rho(\tau)$  [9, 41], rather than by the bandwidth-dependent loss of coherence. However, the model employed in the present study is not expected to be the only model that can account for these data. All previous models that employ relatively narrow filters [10, 29] or a correspondingly steep  $\rho(\tau)$  function [8] are expected to be isomorphic for the present data [43]. Our model is effectively the same as the analytic description and fitting from Rabiner et al. [6], their Equation (13) and, amongst others, conceptually identical to what was proposed by Langford and Jeffress [2]. Following-up on the present proof of concept, the model should be re-implemented to detect IPD fluctuations at arbitrary baseline IPDs. A second comprehensive extension would be a multi-channel approach, ultimately including a binaural interference process. This extension would make the concept compatible to be tested against dozens of datasets from binaural unmasking experiments (see, e.g., [44]). Within its range of validity, the proposed model also offers a computationally more efficient alternative to calculating the running cross-correlation function.

## 5 Conclusions

Combined parametric variation of noise bandwidth and noise ITD revealed that the 2–4 ms noise-ITD range is particularly informative for estimating the filter bandwidth. A simplistic binaural model based on the physiologically plausible concept of extracting interaural phase fluctuations [15] can explain the data based on the monaurally derived auditory-filter bandwidth without requiring delay lines. The results of the present study (i) bridge between the mathematical concept of coherence and auditory feature extraction and (ii) help to resolve the discrepancy between physiologic reports that cast doubt on neural compensation for larger ITDs [45–47] and psychoacoustically motivated models that appeared to require ITD compensation [8–10, 29].

## Data availability statement

The model code is available in the auditory modeling toolbox [48], data on Zenodo under the reference <https://doi.org/10.5281/zenodo.5410778> [49].

## Conflict of interest

The authors declare that they have no conflicts of interest in relation to this article.

## Acknowledgments

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme grant agreement no. 716800 (ERC Starting Grant to Mathias Dietz) and the Deutsche Forschungsgemeinschaft DFG (SFB 1330 - 352015383 - Project B4 to Mathias Dietz and project A2 to Stephan Ewert).

## References

1. L.J. Hirsh: The influence of interaural phase on summation and inhibition. *The Journal of the Acoustical Society of America* 20 (1948) 536–544.
2. T.L. Langford, L.A. Jeffress: Effect of noise crosscorrelation on binaural signal detection. *The Journal of the Acoustical Society of America* 36 (1964) 1455–1458.
3. L.R. Bernstein, C. Trahiotis: Effects of interaural delay, center frequency, and no more than “slight” hearing loss on precision of binaural processing: Empirical data and quantitative modeling. *The Journal of the Acoustical Society of America* 144 (2018) 292–307. <https://doi.org/10.1121/1.5046515>.
4. L.A. Jeffress: A place theory of sound localization. *The Journal of Comparative and Physiological Psychology* 41 (1948) 35–39.
5. C.E. Carr, M. Konishi: A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience* 10 (1990) 3227–3246. <https://doi.org/10.1523/JNEUROSCI.10-10-03227.1990>.
6. L.R. Rabiner, C.L. Laurence, N.I. Durlach: Further Results on Binaural Unmasking and the EC Model. *Journal of the Acoustical Society of America* 40 (1966) 62–70. <https://doi.org/10.1121/1.1910065>.
7. R.M. Stern, G.D. Shear: Lateralization and detection of low-frequency binaural stimuli: effects of distribution of internal delay. *Journal of the Acoustical Society of America* 100 (1996) 2278–2288. <https://doi.org/10.1121/1.417937>.
8. M. van der Heijden, C. Trahiotis: Masking with interaurally delayed stimuli: the use of “internal” delays in binaural detection. *Journal of the Acoustical Society of America* 105 (1999) 388–399.
9. J. Breebaart, S. Van De Par, A. Kohlrausch: Binaural processing model based on contralateral inhibition. I. Model structure. *Journal of the Acoustical Society of America* 110 (2001) 1074–1088.
10. L.R. Bernstein, C. Trahiotis: Binaural detection as a joint function of masker bandwidth, masker interaural correlation, and interaural time delay: Empirical data and modeling”. *Journal of the Acoustical Society of America* 148 (2020) 3481–3488. <https://doi.org/10.1121/10.0002869>.
11. M. Dietz, G. Ashida: Computational models of binaural processing, in *Binaural Hearing*, Litovsky R.Y., Goupell M. J., Fay R.R., Popper A.N., Eds., Springer, New York. 2021, pp. 281–315. <https://doi.org/10.1007/978-3-030-57100-9>.
12. B.R. Glasberg, B.C.J. Moore: Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 47 (1990) 103–138. [https://doi.org/10.1016/0378-5955\(90\)90170-T](https://doi.org/10.1016/0378-5955(90)90170-T).
13. M. Dietz, S.D. Ewert, V. Hohmann: Auditory model based direction estimation of concurrent speakers from binaural signals. *Speech Communication* 53 (2011) 592–605. <https://doi.org/10.1016/j.specom.2010.05.006>.
14. M. Dietz, S.D. Ewert, V. Hohmann, B. Kollmeier: Coding of temporally fluctuating interaural timing disparities in a binaural processing model based on phase differences. *Brain Research* 1220 (2008) 234–245. <https://doi.org/10.1016/j.brainres.2007.09.026>.
15. P.X. Joris, B. van de Sande, A. Recio-Spinoso, M. van der Heijden: Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. *Journal of Neuroscience* 26 (2006) 279–289.
16. R. Bamler, D. Just: Phase statistics and decorrelation in SAR interferograms IGARSS '93, in *Better understanding of earth environment*. IEEE, Tokyo, Japan. 1993, pp. 980–984. <https://doi.org/10.1109/IGARSS.1993.322637>.
17. M.J. Goupell, W.M. Hartmann: Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects. *Journal of the Acoustical Society of America* 119 (2006) 3971–3986. <https://doi.org/10.1121/1.2200147>.
18. M.M. Sondhi, N. Guttman: Width of the spectrum effective in the binaural release of masking. *Journal of the Acoustical Society of America* 40 (1966) 600–606.
19. A.J. Kolarik, J.F. Culling: Measurement of the binaural auditory filter using a detection task. *Journal of the Acoustical Society of America* 127 (2010) 3009–3017. <https://doi.org/10.1121/1.3365314>.
20. S. van de Par, A. Kohlrausch: Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters. *Journal of the Acoustical Society of America* 106 (1999) 1940–1947.
21. J. Breebaart, S. Van De Par, A. Kohlrausch: Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters. *Journal of the Acoustical Society of America* 110 (2001) 1089–1104.
22. T. Marquardt, D. McAlpine: Masking with interaurally “double-delayed” stimuli: The range of internal delays in the human brain. *Journal of the Acoustical Society of America* 126 (2009) EL177–EL182. <https://doi.org/10.1121/1.3253689>.
23. M. Mc Laughlin, J.N. Chabwine, M. van der Heijden, P.X. Joris: Comparison of bandwidths in the inferior colliculus and the auditory nerve. II: Measurement using a temporally manipulated stimulus. *Journal of Neurophysiology* 100 (2008) 2312–2327. <https://doi.org/10.1152/jn.90252.2008>.
24. M. Mc Laughlin, B. van de Sande, M. van der Heijden, P.X. Joris: Comparison of bandwidths in the inferior colliculus and the auditory nerve. I. Measurement using a spectrally manipulated stimulus. *Journal of Neurophysiology* 98 (2007) 2566–2579. <https://doi.org/10.1152/jn.00595.2007>.
25. L.R. Bernstein, C. Trahiotis: Behavioral manifestations of audiometrically-defined “slight” or “hidden” hearing loss revealed by measures of binaural detection. *Journal of the Acoustical Society of America* 140 (2016) 3540–3548. <https://doi.org/10.1121/1.4966113>.
26. S.D. Ewert: AFC – A modular framework for running psychoacoustic experiments and computational perception models, in *Proceedings of the International Conference on Acoustics AIA-DAGA2013*, Merano, Italy. 2013, pp. 1326–1329.
27. H. Levitt: Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America* 49 (1971) 467–477.
28. S. Thavam, M. Dietz: Smallest perceivable interaural time differences. *Journal of the Acoustical Society of America* 145 (2019) 458–468.
29. L.R. Bernstein, C. Trahiotis: Converging measures of binaural detection yield estimates of precision of coding of interaural temporal disparities. *Journal of the Acoustical Society of America* 141 (2017) 1150–1160. <https://doi.org/10.1121/1.4935606>.
30. W.T. Bourbon, L.A. Jeffress: Effect of bandwidth of masking noise on the detection of homophasic and antiphase tonal signals. *Journal of the Acoustical Society of America* 37 (1965) 1180–1181.



31. V. Hohmann: Frequency analysis and synthesis using a Gammatone filterbank. *Acta Acustica United with Acustica* 88 (2002) 433–442.
32. M.W.H. Remme, R. Donato, J. Mikiel-Hunter, J.A. Ballesteros, S. Foster, J. Rinzel, D. McAlpine: Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues. *Proceedings of the National Academy of Sciences of the United States of America* 111 (2014) E2339–E2348. <https://doi.org/10.1073/pnas.1316216111>.
33. N.I. Durlach: Binaural signal detection: Equalization and cancellation theory, in *Foundations of Modern Auditory Theory*, Tobias J., Ed., Vol. 2, Academic Press, New York, 1972, pp. 369–462.
34. S.D. Ewert, N. Parouty, C. Lorenzi: A two-path model of auditory modulation detection using temporal fine structure and envelope cues. *European Journal of Neuroscience* 51 (2020) 1265–1278. <https://doi.org/10.1111/ejn.13846>.
35. E. Zwicker, G.B. Henning: The four factors leading to binaural masking-level differences. *Hearing Research* 19 (1985) 29–47. [https://doi.org/10.1016/0378-5955\(85\)90096-6](https://doi.org/10.1016/0378-5955(85)90096-6).
36. S.K. Isabelle, H.S. Colburn: Detection of tones in reproducible narrow-band noise. *Journal of the Acoustical Society of America* 89 (1991) 352–359. <https://doi.org/10.1121/1.400470>.
37. M. van der Heijden, P.X. Joris: Interaural correlation fails to account for detection in a classic binaural task: Dynamic ITDs dominate N0S $\pi$  detection. *Journal of the Association for Research in Otolaryngology* 11 (2010) 113–131. <https://doi.org/10.1007/s10162-009-0185-8>.
38. H. Lüddemann, H. Riedel, B. Kollmeier: Logarithmic scaling of interaural cross correlation: a model based on evidence from psychophysics and EEG, in *Hearing: From sensory processing to perception – 14th International Symposium on Hearing*, Kollmeier B, Klump G, Hohmann V, Langemann U, Mauermann M, Uppenkamp S, Verhey J, Eds., Springer, Berlin, 2007, pp. 379–388.
39. J.F. Culling, H.S. Colburn, M. Spurchise: Interaural correlation sensitivity. *Journal of the Acoustical Society of America* 110 (2001) 1020–1029. <https://doi.org/10.1121/1.1383296>.
40. M. Dietz, J.H. Lestang, P. Majdak, R.M. Stern, T. Marquardt, S.D. Ewert, W.M. Hartmann, D.F. Goodman: A framework for testing and comparing binaural models. *Hearing Research* 360 (2018) 92–106.
41. C. Trahiotis, L.R. Bernstein, M.A. Akeroyd: Manipulating the “straightness” and “curvature” of patterns of interaural cross correlation affects listeners’ sensitivity to changes in interaural delay. *Journal of the Acoustical Society of America* 109 (2001) 321–330.
42. C.A. Shera, J.J. Guinan, A.J. Oxenham: Otoacoustic estimation of cochlear tuning: Validation in the chinchilla. *Journal of the Association for Research in Otolaryngology* 11 (2010) 343–365. <https://doi.org/10.1007/s10162-010-0217-4>.
43. R.H. Domnitz, H.S. Colburn: Analysis of binaural detection models for dependence on interaural target parameters. *Journal of the Acoustical Society of America* 59 (1976) 598–601.
44. J.F. Culling, M. Lavandier: Binaural and spatial unmasking, in *Binaural Hearing*, Litovsky R.Y., Goupell M.J., Fay R.R., Popper A.N., Eds., Springer, New York, 2021, pp. 209–242. <https://doi.org/10.1007/978-3-030-57100-9>.
45. P.X. Joris, B. van de Sande, D.H. Louage, M. van der Heijden: Binaural and cochlear disparities. *Proceedings of the National Academy of Sciences of the United States of America* 103 (2006) 12917–12922.
46. D. McAlpine, D. Jiang, A.R. Palmer: A neural code for low-frequency sound localization in mammals. *Nature Neuroscience* 4 (2001) 396–401.
47. S.K. Thompson, K. von Kriegstein, A. Deane-Pratt, T. Marquardt, R. Deichmann, T.D. Griffiths, D. McAlpine: Representation of interaural time delay in the human auditory midbrain. *Nature Neuroscience* 9 (2006) 1096–1098.
48. P. Majdak, C. Hollomey, R. Baumgartner: AMT 1.0: The toolbox for reproducible research in auditory modeling. Submitted to *Acta Acustica*. [Google Scholar]
49. M. Dietz, J. Encke, K.I. Bracklo, S.D. Ewert: Data for the Article: Prediction of tone detection thresholds in interaurally delayed noise based on interaural phase difference fluctuations (0.2). Zenodo, 2021. <https://doi.org/10.5281/zenodo.5410778>.

**Cite this article as:** Dietz M, Encke J, Bracklo KI, & Ewert SD. 2021. Tone detection thresholds in interaurally delayed noise of different bandwidths. *Acta Acustica*, 5, 60.