VI

Publication VI

Hirvonen, T. and Pulkki, V., ”Predicting Binaural Masking Level Difference and Dichotic Pitch Using Instantaneous ILD Model”, AES 30th Int. Conference, 2007.

Predicting Binaural Masking Level Difference and Dichotic Pitch Using Instantaneous ILD Model

Toni Hirvonen[1] and Ville Pulkki[1]

[1] *Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Finland*

Correspondence should be addressed to Toni Hirvonen (`Toni.Hirvonen@tkk.fi`)

ABSTRACT

A recently proposed auditory model is examined using simulated binaural masking level difference (BMLD) and dichotic pitch (Huggins and binaural edge pitch) stimuli. The model is based on calculating the instantaneous interaural level difference, i.e. the difference between the left and right ear neural signals. The model output produces pronounced maximum at the signal frequency with BMLD stimulus. Both dichotic pitch stimuli produce a notable maxima at the pitch frequencies. Although the model can thus be interpreted to predict known psychoacoustical results, an exact quantitative comparison with the model responses and data from BMLD and binaural pitch experiments is not performed. Rather, this paper serves as a "proof of concept".

1. INTRODUCTION

Auditory modeling refers to the use of computational models to predict the human responses to auditory stimuli. The obvious criterion of a perceptual model is how well it can be used for the specific task at hand. However, in order for a model to be practical, it should not be overly complex and also be generalizable to different situations. Additionally, auditory modeling research aims to a better understanding of human physiology by mimicking its known functions.

In this paper, an auditory model that was inspired by recent neurophysiological results and theorems is introduced and utilized. The model was originally implemented to characterize the perception of interaural coherence and the interaural level difference localization cue [1]. The research question of this work is whether the model can also be applied to account for two psychoacoustical phenomena: 1) binaural masking level difference (BMLD), and 2) dichotic pitch. These can both be thought as different forms of binaural detection, i.e. the effects do not arise with monaural listening.

The following subsections briefly review these phenomena and the relevant research. Commonly used auditory modeling techniques are also discussed. Section 2 illustrates the implementation of the present model and discusses its physiological relevance. In Sections 3 and 4, the model simulation results for BMLD and dichotic pitch stimuli are presented. Finally, additional discussion and a summary are given in Sections 5 and 6.

1.1. Binaural Masking Level Difference

Auditory masking level is derived from the threshold where a target signal is detected amongst a masker signal. Starting with the research by Hirsh [2], it has been established that if different interaural manipulations are applied to the signal and masker, the detection threshold is likely to be reduced. As such binaural tests were easy to perform with early auditory equipment, namely with headphones, extensive data exists on the subject. The most known classic test paradigm is to present a signal tone that is phase-shifted between the ears by 180° among a wideband masker noise (N_0S_π), and compare the obtained threshold to the N_mS_m (monaural signal and masker) or to the N_0S_0 (diotic signal and masker) reference.

In general, the classic BMLD is prominent (up to 15 dB) at low signal frequencies, but also present (up to 5 dB) at the higher frequency range above approximately 2 kHz. When presenting the signal monaurally (N_0S_m), BMLD is approximately 6 dB smaller than in the N_0S_π configuration. However, BMLD may occur in any situation where the signal and the masker have different interaural parameters. Many different test paradigms on the topic have been conceived, see e.g. [3] for a review on earlier research and [4] for more recent experiments.

Cross-correlation auditory models, based on the Jeffress coincidence-counting concept [5], have been widely utilized to predict simple psychoacoustical data such as classic BMLD experiments [6]. The patterns of coincidence-counting activity show "dimples" caused by the S_π signal among the masker. A common approach is to assume that the activity pattern is analyzed by an upper-level pattern recognition system, and thus the pattern can be interpreted visually. A more complicated method is to apply statistical methods and try to fit the model output to psychoacoustical results. This can be difficult especially if the listening task is complex. Colburn was able to describe much of the classical BMLD results by forming a decision (signal present or not) variable based on the coincidence output [7]. In addition to coincidence-counting models, equalization-cancellation (EC) model that is based on the differences rather than similarities of the ear signals has long been utilized for predicting BMLDs [8]. The basic idea of the model is to cancel out the masker and leave the interaurally varying component, i.e. the signal for detection.

There are BMLD situations that are problematic to the presently available auditory models. Colburn et al. performed experiments using frozen, or reproducible noise samples as maskers [9] [10]. It was found that the detection performance varies between individual masker waveforms and this has proven to be a difficulty for the traditional models. Recently, Davidson et al. compared the use of specific BMLD models based on stimulus energy and on temporal stimulus structure [11]. The best predictions were obtained with a model based on a weighted combination of energy in multiple critical bands. However, they argue that energy-based models are unable to predict all types of stimuli and that temporal models that emphasizes waveform differences should also be considered in the future.

1.2. Dichotic Pitch

Dichotic pitch occurs when two different broadband ear signals induce a pitch perception with simultaneous presentation but fail to do so when the two signals are presented monaurally. In a sense, dichotic pitches fall into the category of auditory illusions as there is no real signal to detect. Huggins pitch is the first and the most famous of the dichotic pitches [12]. The corresponding stimulus is implemented by creating a phase transition of 360° during a narrow frequency band of a broadband noise. This paper also considers binaural edge pitch, whose corresponding stimulus is similar to Huggins pitch stimulus, other than the phase transition being 180°. The percept that occurs with the previous two stimuli clearly has a specific frequency similarly as the percepts created by a sinusoid or narrowband noise. There are also dichotic pitch stimuli which are perceived more of a complex-tone-like but they are not examined in this paper.

Auditory models can be utilized for predicting the perception of dichotic pitch similarly as in the BMLD case [13] [6]. Cross-correlation models produce a correlation peak at the pitch frequency, whereas EC models produce a decorrelation peak. Culling et. al. have compared the two approaches for various dichotic pitches [14] [15]. They concluded that a modified EC model provides the most coherent prediction results in the tested cases. The model investigated in this paper bears some resemblance to the EC model principles, as explained in the following section.

2. MODEL IMPLEMENTATION

2.1. Basis for the Model

After the sound travels through the outer and middle parts of the ear, the acoustic signals are transformed into neural impulses in the Cochlear Nucleus at separate critical frequency bands. Superior Olivary Complex (SOC) is one of the first sites of binaural interaction [16]. SOC contains two smaller organs that are important in binaural hearing: Medial Superior Olive (MSO) and Lateral Superior Olive (LSO). The exact neural mechanisms of these organs are not known, but the consensus is that MSO and LSO are predominantly responsible for the encoding of the two primary interaural cues: interaural time difference (ITD) and level difference (ILD), respectively.

The authors are currently developing a general auditory model intended to predict common hearing phenomena including the perception of ITD and ILD. This project was inspired by recent results that partially question the physiological validity of the existing modeling methods [17] [18]. The general model is a composite of different parts similarly to the actual hearing system. The second author has developed an ITD model that contributes strongly to for example low-frequency localization [19]. The model presented in this paper is designed to estimate only the ILD processing. In this sense, these two parts would process the signal analogously to MSO and LSO.

However, a method of combining the information from these two models is still under development.

The fundamental idea of the ILD model is to investigate the instantaneous difference between the ear signals, rather than consider the ILD as a long-time level difference. A number of studies have shown that the ITD mechanism is quite sluggish, with a time constant approximately between 100 and 250 ms depending on the listening task, and that ILD decoding is much faster [20] [18]. Joris and Yin have found that a fast subtractive mechanism could facilitate the sensitivity to the ITDs of signal envelopes in the LSO [17]. These findings indicate that the LSO might also contribute to processes that have been traditionally thought to be time-based in nature.

The present authors have previously utilized the model for predicting the perceived ILD localization cue and interaural coherence [1]. The results corresponded roughly to psychoacoustical data and recent results have also suggested that the interaural coherence detection is based on the short-time temporal fluctuations of the signal [21]. This research is a natural continuum for the previous coherence experiments, as there is evidence that binaural detection and interaural coherence discrimination are effected by a common perceptual mechanism [22] [23].

2.2. Model Structure

The model implementation used in this paper is illustrated in Figure 1. Generally, the signals ascend from left to right, as indicated by arrows. The input signals to the cochleas of the left and right ear can be filtered with appropriate head-related transfer functions (HRTFs) to simulate the outer ear prior to the model, depending on the application. In these experiments the use of HRTFs is omitted due to the fact that the corresponding psychoacoustical experiments have been done with headphones.

The first stage of processing involves a gammatone filterbank (GTFB) for both ear signals that divides the signal into critical frequency bands similarly as in the cochlea [24]. The equivalent rectangular bandwidth (ERB) scale [25] was used such that the resolution was one filter per ERB-band. These bands are processed separately from this point on, i.e. Figure 1 shows the processing for only one critical band. No across-channel effects are considered in the present paper. There is some evidence of across-critical band processing occurring in the auditory system (e.g. [26]) but combining the information of different bands is likely to occur at the higher levels of the auditory system.

After the gammatone filterbank, the signals are transformed into neural form. The model is probabilistic in the sense that individual neural spikes are not modeled. Rather, the most simple neural transform is half-wave rectification as neural impulses cannot be negative, and it is also used in this paper. More physiologically plausible transforms were also experimented with but the results were not affected significantly. A neural lowpass filter is also applied to the signals in order to simulate the loss of waveform synchronization at high frequencies (approx. higher than 1000-2000 Hz). The neural lowpass filter was realized using a 4. degree IIR with a cutoff frequency of 700 Hz (i.e. $\tau \approx 0.23$ ms). This filter has a similar magnitude response as the neural filter derived by Bernstein and Trahiotis [23]. Next, internal Gaussian noise normalized to 0 dB SPL is added to both ear signals. The normalization coefficient for the desired SPL level is in this paper calculated from the signal RMS-value using the rule:

$$\text{RMS} = 1 \Rightarrow \text{SPL} = 0 \text{ dB}$$

The parts so far constitute the so called peripheral processing of the model.

The next stages beyond the periphery implement the ILD-part of binaural processing. The model presented here is not intended to be exactly accurate physiologically, as the structure of the LSO is not completely known. However, most of the cells in LSO have been shown to be inhibition-excitation (IE) type [18]. IE refers to a process where the contralateral ear signal inhibits the ipsilateral input. For this reason, the ILD processing is in the model approximated with a simple subtraction: the contralateral ear signal is subtracted from the ipsilateral signal sample-by-sample to calculate the instantaneous ILD for the left and right sides, whose negative parts are then removed with half-wave rectification. This process is similar to the process hypothesized by Joris and Yin that facilitates the interaural phase difference sensitivity of LSO cells [17].

If the ear signals are the same, the reduction yields zero output. However, differing signals leave parts of the original ipsilateral signal intact. The basic principle of this mechanism is illustrated in Figure 2. The following lowpass filter (1. degree IIR, τ =5 ms) functions as a temporal integrator and simulates the slowness and saturation of the neural cells. After the half wave rectification, both

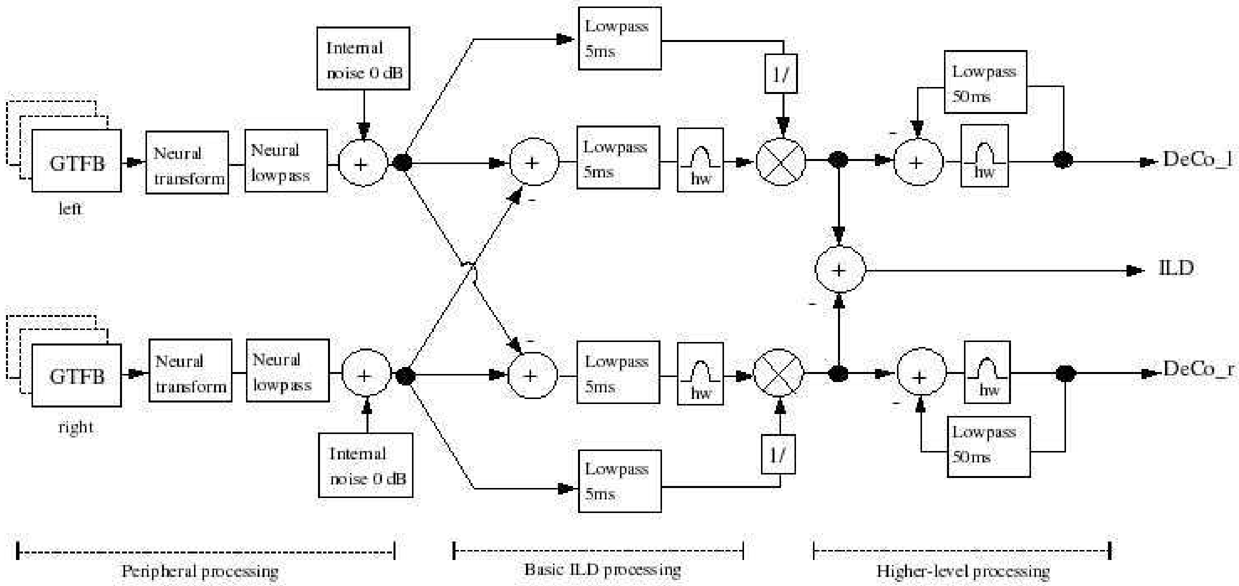

Fig. 1: Model implementation used for the simulations in this research. See text for description.

ear signals are divided sample-by sample with the ipsilateral signal filtered with the same filter as used in the temporal integrator. This is done for sake of normalizing the outputs as relative to the input and between frequency channels.

After the binaural interaction, the signal information is used to extract psychoacoustically relevant information as in the higher stages of auditory processing. There are two specific outputs: the ILD-channel, and decorrelation-channels (DeCo). In previous experiments, the DeCo-channels were used to estimate the perceived coherence cue, whereas the ILD-channel predicted the ILD localization cue [1]. In this paper, solely DeCo-outputs are used. They were designed to increase as the correlation between the ear signals decreases. The processing includes a negative feedback loop with a low-pass filter (1. degree IIR, τ =50 ms). This mechanism is used to remove the steady DC-component of the DeCo-output; in an anechoic environment, a sound coming from non-zero azimuth causes a constant, non-varying ILD, which in the model manifests itself as DC-component in the DeCo-channel. Thus it is appropriate to remove this, focusing only on the time-varying component.

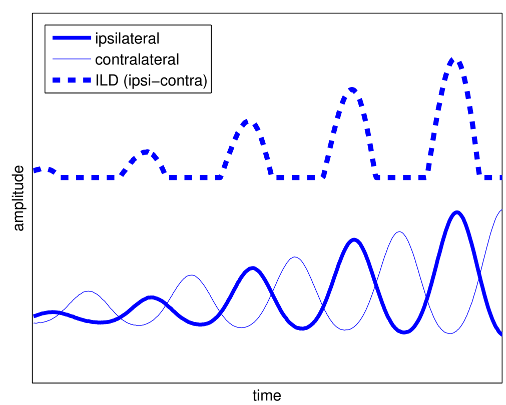

Fig. 2: Illustration of the principle of the instantaneous ILD calculation for the left ear signal. The contralateral channel is subtracted from the ipsilateral signal and half-wave rectified. The calculation result (ILD) has been shifted in amplitude to clarify the figure.

2.3. Relation to other models

The EC model proposed by Durlach [8] bears resemblance to the present model in the sense that it is also based on the subtraction of the two ear signals. However, the original EC-model includes specific interaural phase shift and level compensation before signal subtraction. Recently, Breebaart et. al. have suggested and extensively examined a general-purpose auditory model that is based on EC-type elements [27]. Their model does not include interaural phase or level manipulation but rather the EC-elements are located in a 2-dimensional delay and attenuation network that produces an activation pattern.

The present model differs from the previous EC-type implementations in that it includes filters as well as signal division and addition in order to extract the desired effects from the original inputs. However, the most important difference is probably in the design philosophy; the present model aims specifically to implement the ILD processing-part of the hearing that occurs mainly in the LSO. The ITD cue is to be calculated with a separate, MSO-inspired model.

3. MODELING BINAURAL MASKING LEVEL DIFFERENCE

This section presents the model simulation results for some common BMLD cases. Direct comparisons with psychoacoustical data are not performed in this paper. Rather, it is demonstrated how the model mechanism itself can produce output activity that could be used by the higher-level pattern analysis similarly as in [6].

Figure 3 illustrates the model output for a typical BMLD case: a signal (500 Hz tone) is masked by a wideband Gaussian white noise (N_0S_π). Signal-to-noise ratio (SNR) was -10 dB, which is slightly above the audible threshold for a 500 Hz-signal. Additionally, outputs for the signal (S_π) and the diotic masker alone (N_0), as well as for a monaural signal with masker (N_0S_m, SNR=-10 dB) are shown. All four stimuli were normalized to a 70 dB SPL relative to the 0 dB internal noise. Both left and right outputs were averaged over the 1 s stimulus length (L in samples) for individual critical bands (center frequencies 100 - 8190 Hz) and summed to yield the mean output:

$$\text{mean out}(i) = 20log_{10}\left(\frac{1}{L}\sum_{L} DeCo_l(i) + \frac{1}{L}\sum_{L} DeCo_r(i)\right)$$

where index i indicates a given critical frequency band.

The mean output can be examined as a function of frequency. It can be seen that the N_0S_π case output peaks at the signal frequency because at that frequency region, the left and the right signals are different. This illustrates how the model can be used to interpret the frequency of the signal from a typical BMLD stimulus. When the diotic masker noise alone is presented, output is very low, since both ear signals are similar.

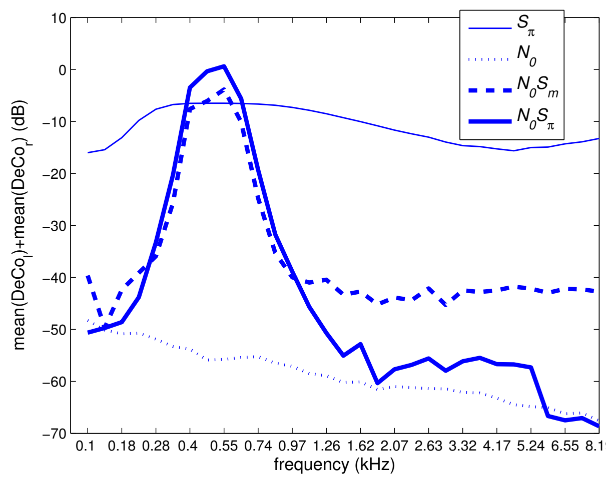

Fig. 3: Model output calculated from the DeCo-channel time averages with a typical BMLD N_0S_π stimulus. Masker was Gaussian white noise and the signal frequency 500 Hz with an SNR of -10 dB. Additionally, outputs for the signal and the masker alone, as well as for a monaural signal with masker (N_0S_m) are shown. All four stimuli were normalized for input level.

An interesting result is seen from the output for the dichotic signal presented alone (S_π); there is no clear peak at the signal frequency and the overall maximum is lower than in N_0S_π and N_0S_m cases. This indicates that the masker is required for the output peak to appear, which is naturally consistent with psychoacoustics where a masker is required for the BMLD itself to occur. The sample-by sample division occurring after the binaural interaction (see Figure 1) flattens the output peak when a mere dichotic tone is presented.

It should be noted that a N_0S_0 case would produce a similar low output as the diotic noise alone. Modeling monaural signal detection among the masker would require the addition of monaural paths to the model, which have not yet been implemented. For this reason, the traditional BMLD comparison between the N_0S_0 and N_0S_π cases is not possible here. Furthermore, calculating BMLD this way would require forming a decision variable for the threshold of signal detection.

In psychoacoustical experiments, the N_0S_m case has been shown to produce up to 6 dB smaller BMLDs than the traditional N_0S_π case. With the signal frequency of 500 Hz and SNR of -10 dB, the signal is near the detection threshold in the N_0S_m case. The results in Figure 3 indicate that the N_0S_π configuration produces a larger output with these parameters. The model thus gives results similar to psychoacoustical data in this respect.

As mentioned in Section 2, the variance of BMLD between individual masker waveforms has proved problematic for many auditory models. Figure 4 presents the model output calculated similarly as in Figure 3 for 15 random N_0S_π cases. The signal frequency was again 500 Hz and SNR -10 dB. However, the masker was narrowband Gaussian noise (445-561 Hz) and the stimulus length was only 300 ms. These parameters correspond to those used by Isebelle and Colburn in psychoacoustical experiments [10]. The results indicate that the output level varies at the signal frequency. This can be taken as an indicator of the models capacity to achieve differences between individual maskers. The differences seem relatively small considering that listening tests have given drastic masking differences. However, it should be noted that the output was simply averaged over the entire stimulus length (300 ms) and more complex methods may be tried in the future.

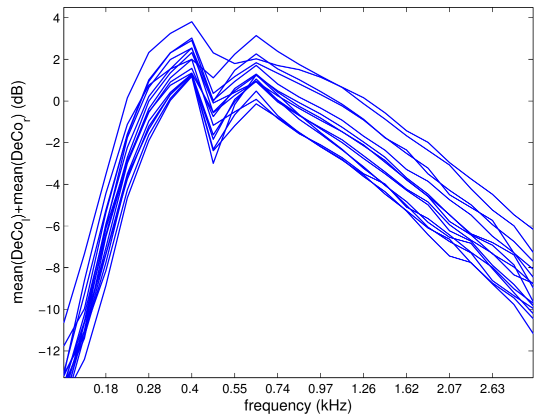

Fig. 4: The model output for 15 N_0S_π random individual masker cases. The signal frequency was 500 Hz and SNR -10 dB with a stimulus length of 300 ms. The masker was narrowband Gaussian noise (445-561 Hz) .

4. MODELING BINAURAL PITCH

Analogously as with the BMLD stimuli, this section presents the model simulations for the two dichotic pitch cases that were discussed in Section 1.2. The length of each stimulus was 1 s, and the input stimuli were normalized to 70 dB RMS level relative to the internal noise level. The model output was again calculated as a time average of the DeCo-channel signals similarly as in the previous section. The results are presented in Figure 5.

The parameters for Huggins pitch and binaural edge pitch stimuli were similar, other than the amount of phase transition, which is by definition 360° for the former and 180° for the latter stimulus. The noise bandwidth was 1000 Hz (100-1100 Hz), with the transition bandwidth being 50 Hz (475-525 Hz). It can be seen from Figure 5 that the output produces a pronounced peak value at the frequency region around 500 Hz where the pitches in both cases are also perceived subjectively. Thus we conclude that HP and BEP phenomena could be mediated by similar mechanisms as implemented in the model.

5. DISCUSSION

Auditory modeling research aims to examine if the produced model corresponds to the actual perceptions of humans. Within the scope of this paper, the aim was to show that the proposed mechanism is able to produce advantage in binaural signal detection tasks. If direct comparison between the model data and psychoacoustical results is required, a common approach is to form a decision variable. However, one viewpoint is that the model output should produce an effect that is clearly seen without the use of statistics. This is the approach taken in this paper.

The results in this paper in both BMLD and dichotic pitch cases are presented using the signal frequency of 500 Hz. The similar stimuli with other target frequencies were experimented with and they produced similar results. Detailed comparison of different signal frequencies is left to future studies.

It can be argued that dichotic pitch does not fall into the

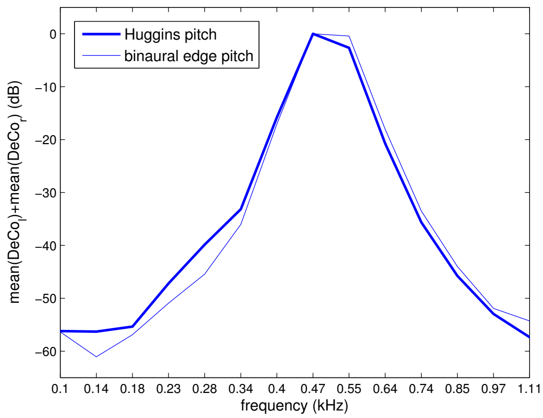

Fig. 5: The model output for Huggins pitch and binaural edge pitch stimuli. The noise bandwidth was 100-1100 Hz, and the phase transition bandwidth where the dichotic pitch is heard was 475-525 Hz in both cases.

signal detection category as such because there is no actual target signal, and thus the phenomenon is not comparable to BMLD. A question for future research is that could dichotic pitch phenomena mediated by the binaural perception of timbre as the "pitch" itself is often perceived more like the pitch produced by a narrowband noise than the pitch of a pure tone.

6. SUMMARY

A novel auditory model implementation was examined with simulated BMLD and dichotic pitch cases. The model is based on calculating the instantaneous ILD, i.e. the difference of neural signals between the ears. The model implements the ILD processing part of the binaural system, and should be complemented by mechanisms that implement the ITD, and monaural processing of the hearing system.

However, some effects were seen using the ILD model alone. The model output produced pronounced peaks at the signal frequencies with BMLD stimulus. The dichotic pitch cases, Huggins pitch and binaural edge pitch stimuli produced a notable peak at the pitch frequency. No direct comparison with psychoacoustical data is given is this paper. Rather, the model output is thought to be processed by upper-level pattern recognition, or similar systems.

7. ACKNOWLEDGEMENTS

This work was supported by The Academy of Finland (projects no. 201050 and 105780) and by the Emil Aaltonen Foundation.

8. REFERENCES

[1] T. Hirvonen and V. Pulkki, "Interaural coherence estimation with instantaneous ILD," in *Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG 2006)*. Reykjavik, Iceland, June 2006.

[2] I. J. Hirsh, "The influence of interaural phase on interaural summation and inhibition," *J. Acoust. Soc. Am.*, vol. 20, no. 4, pp. 536–544, 1948.

[3] N. I. Durlach and H. S. Colburn, "Binaural phenomenna," in *Handbook of Perception*, E. C. Carterette and M. P. Friedman, Eds., pp. 365–466. Academic Press, 1978.

[4] R. H. Gilkey and T. R. Anderson, Eds., *Binaural and Spatial Hearing in Real and Virtual Environments*, Lawrence Erlbaum Assoc., Mahwah, NJ, US, 1997.

[5] L. A. Jeffress, "A place theory of sound localization," *J. Comp. Physiol. Psych.*, vol. 41, pp. 35–39, 1948.

[6] R. M. Stern and C. Trahiotis, "Models of binaural perception," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds., pp. 499–531. Lawrence Erlbaum assoc., Mahwah, New Jersey, 1997.

[7] H. S. Colburn, "Theory of binaural interaction based on auditory-nerve data. II. detection of tones in noise," *J. Audio Eng. Soc.*, vol. 61, no. 2, pp. 525–533, 1977.

[8] N.I. Durlach, "Equalization and cancellation theory of binaural masking-level differences," *J. Acoust. Soc. Am.*, vol. 35, no. 8, pp. 1206–1218, 1963.

[9] H. S. Colburn, S. K. Isabelle, and D. J. Tollin, "Modeling binaural detection performance for individual waveforms," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds., pp. 533–555. Lawrence Erlbaum assoc., Mahwah, New Jersey, 1997.

[10] S. K. Isabelle and H. S. Colburn, "Detection of tones in reproducible narrow-band noise," *J. Acoust. Soc. Am.*, vol. 89, no. 1, pp. 352–359, 1991.

[11] S. A. Davidson, R. H. Gilkey, H. S. Colburn, and L. H. Carney, "Binaural detection with narrow-band and wideband reproducible noise maskers. III. monaural and diotic detection and model results," *J. Acoust. Soc. Am.*, vol. 119, no. 4, pp. 2258–2275, 2006.

[12] E. M. Cramer and W. H. Huggins, "Creation of pitch through binaural interaction," *J. Acoust. Soc. Am.*, vol. 30, no. 5, pp. 413–417, 1958.

[13] P. Jeffrey Bloom, "Pitch of noise signals: Evidence for a "central spectrum"," *J. Audio Eng. Soc.*, vol. 61, no. 1, pp. 150–161, 1977.

[14] J. F. Culling, A. Q. Summerfield, and D. H. Marshall, "Dichotic pitches as illusions of binaural unmasking. I. huggins' pitch and the binaural edge pitch," 1998.

[15] J. F. Culling, D. H. Marshall, and A. Q. Summerfield, "Dichotic pitches as illusions of binaural unmasking. II. the fourcin pitch and the dichotic repetition pitch," *J. Acoust. Soc. Am.*, vol. 103, no. 6, pp. 3527–3539, 1998.

[16] W. A. Stotler, "An experimental study of the cells and connections of the superior olivary complex of a cat," *J. Comp. Neurol.*, vol. 100, pp. 401–423, 1953.

[17] P. X. Joris and T. C. Yin, "Envelope coding in the lateral superior olive. I. sensitivity to interaural time differences," *J. Neurophysiol.*, vol. 73, no. 3, pp. 1043–1062, 1995.

[18] B. Grothe, "Sensory systems: New roles for synaptic inhibition in sound localization," *Nature Neuroscience*, vol. 4, pp. 540–550, 2003.

[19] V. Pulkki, "Computational two-channel itd model," Paper presented at the AES 120th Convention Paris, France, 2006.

[20] D. W. Grantham, "Discrimination of dynamic interaural intensity differences," *J. Acoust. Soc. Am.*, vol. 76, no. 1, pp. 71–76, 1984.

[21] M. J. Goupell and W. M. Hartmann, "Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects," *J. Acoust. Soc. Am.*, vol. 119, no. 6, pp. 3971–3986, 2006.

[22] M. Jain, D. T. Gallagher, J. Koehnke, and H. S. Colburn, "Fringed correlation discrimination and binaural detection," *J. Acoust. Soc. Am.*, vol. 90, no. 4, pp. 1918–1926, 1991.

[23] L. R. Bernstein and C. Trahiotis, "The normalized correlation: Account for binaural detection across center frequency," *J. Acoust. Soc. Am*, vol. 100, no. 6, pp. 3774–3784, 1996.

[24] M. Slaney, "An efficient implementation of the Patterson-Holdsworth filter bank," Tech. Rep. 35, Apple Computer, Inc., 1993.

[25] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, vol. 47, no. 1-2, pp. 103–138, 1990.

[26] W. A. Yost and S. Sheft, "Across-critical-band processing of amplitude-modulated tones," *J. Acoust. Soc. Am.*, vol. 85, no. 2, pp. 848–857, February 1989.

[27] J. Breebaart, S. van de Par, and A. Kohlrauch, "Binaural processing model based on contralateral inhibition I: Model structure," *J. Acoust. Soc. Am.*, vol. 110, no. 2, 2001.