Binaural processing of modulated interaural level differences

Eric R. Thompson[a] and Torsten Dau[b]
Centre for Applied Hearing Research, Acoustic Technology, Ørsted.DTU, Technical University of Denmark, Building 352, Ørsteds Plads, 2800 Kgs. Lyngby, Denmark

(Received 13 September 2006; accepted 9 November 2007)

Two experiments are presented that measure the acuity of binaural processing of modulated interaural level differences (ILDs) using psychoacoustic methods. In both experiments, dynamic ILDs were created by imposing an interaurally antiphasic sinusoidal amplitude modulation (AM) signal on high-frequency carriers, which were presented over headphones. In the first experiment, the sensitivity to dynamic ILDs was measured as a function of the modulation frequency using puretone, and interaurally correlated and uncorrelated narrow-band noise carriers. The intrinsic interaural level fluctuations of the uncorrelated noise carriers raised the ILD modulation detection thresholds with respect to the pure-tone carriers. The diotic fluctuations of the correlated noise carriers also caused a small increase in the thresholds over the pure-tone carriers, particularly with low ILD modulation frequencies. The second experiment investigated the modulation frequency selectivity in dynamic ILD processing by imposing an interaurally uncorrelated bandpass noise AM masker in series with the interaurally antiphasic AM signal on a pure-tone carrier. By varying the masker center frequencies relative to the signal modulation frequency, broadly tuned, bandpass-shaped patterns were obtained. Simulations with an existing binaural model show that a low-pass filter to limit the binaural temporal resolution is not sufficient to predict the results of the experiments. © *2008 Acoustical Society of America.* [DOI: 10.1121/1.2821800]

PACS number(s): 43.66.Pn, 43.66.Mk [RLF] Pages: 1017–1029

I. INTRODUCTION

Information in sound signals is carried not only by the fine structure of the sound, but also by the intensity fluctuations of its envelope. In a reverberant environment, reflections can reduce the depth of those envelope fluctuations and can change their phase. The effective amount of envelope modulation and modulation phase transmitted to a receiver can be derived from the source–receiver impulse response as a function of the modulation frequency (Schroeder, 1981). This complex modulation transfer function (MTF) shows the modulation attenuation and phase shift as a function of modulation frequency for the particular source–receiver transmission path. A normal human auditory system has two working ears, thereby receiving information from a given source via two transmission paths and through two MTFs. Interaural differences in the modulation phase and/or depth can create fluctuating interaural level differences (ILDs) and interaural time differences (ITDs). In order to understand how ILD fluctuations are perceived and to begin to understand the binaural processing of envelopes in reverberation, artificial stimuli were generated in the present study with sinusoidal amplitude modulation and a controlled interaural modulation phase difference. The stimuli were presented over headphones to listeners in psychoacoustic tests.

An ILD is usually perceived as a lateralization of the sound source toward the ear with the higher intensity sound. When an ILD changes slowly, the sound is perceived to move, while more rapid ILD fluctuations are usually perceived as a stationary sound source with a broad or diffuse sound image (e.g., Blauert, 1972; Grantham, 1984; Griesinger, 1997). This is analogous to the ability of the auditory system to follow slow intensity fluctuations in monaural or diotic stimuli, and the perception of roughness with more rapid fluctuations (e.g., Terhardt, 1968).

Dynamic ILDs can be created by imposing amplitude modulation with an interaural modulation phase difference. However, a static interaural modulation phase difference can also be interpreted as a static envelope ITD, corresponding to the phase difference divided by the angular modulation frequency. High-frequency sounds, which cannot be lateralized based on the ITD of their fine structure (e.g.,Klumpp and Eady, 1956; Mills, 1960), can be lateralized based on the ITD of their envelopes (e.g., Klumpp and Eady, 1956; Henning, 1974; Nuetzel and Hafter, 1981; Bernstein and Trahiotis, 1994). A static modulation phase difference could then create a percept of a sound lateralized toward the leading ear instead of creating a moving or diffuse sound image, depending on the envelope ITD. However, with a phase difference of π, as was used in the present study, it is unclear which ear should be leading because of the temporal symmetry of the sinusoid. For complex sounds with random interaural level fluctuations, those fluctuations may actually be encoded internally as a combination of time-varying ITDs and ILDs. In situations with ambiguous localization cues, such as with a π interaural phase difference, onset cues may dominate localization of the ongoing signal (Buell *et al.*, 1991).

The temporal acuity of the auditory system is often measured by determining the threshold of detection of the sinusoidal modulation of a physical parameter as a function of the modulation frequency, referred to as a "temporal modu-

[a]Electronic mail: et@oersted.dtu.dk.
[b]Electronic mail: tda@oersted.dtu.dk.

lation transfer function" (TMTF). For example, the TMTF with diotic amplitude modulation (AM) was measured by Viemeister (1979) with broadband noise carriers, by Fleischer (1982) and Dau *et al.* (1997a) with narrow-band noise carriers, and by Kohlrausch *et al.* (2000) with pure-tone carriers. Other studies have investigated the temporal acuity to dynamic interaural parameters, such as interaural time or phase differences (e.g., Grantham and Wightman, 1978; Witton *et al.*, 2000) and interaural correlation (e.g., Grantham, 1982). Grantham (1984), Grantham and Bacon (1991), and Stellmack *et al.* (2005) measured the acuity of the binaural system in the detection of modulated ILDs, generated with interaurally antiphasic sinusoidal AM signals.

The TMTFs measured with pure-tone and diotic broadband noise carriers show a high sensitivity to slow AM, with a minimum modulation depth, m, required for detection of around 0.04 (often discussed on a decibel scale as $20\log_{10}m$, here −28 dB) for low frequency modulations. As the modulation rate increases, larger modulation depths are required for detection, thereby exhibiting an overall low-pass characteristic (e.g., Viemeister, 1979; Kohlrausch *et al.*, 2000). However, the thresholds measured with narrow-band noise carriers can exhibit high-pass as well as low-pass characteristics, depending on the bandwidth of the noise (Fleischer, 1982; Dau *et al.*, 1997a). This led Dau and colleagues to propose a modulation filterbank model with bandpass filters acting on the envelope of a stimulus (Dau *et al.*, 1997a, b), which can simulate AM detection performance with narrow-band as well as broadband noise. Bacon and Grantham (1989), Houtgast (1989), and Ewert *et al.* (2002) made more direct measurements of modulation frequency selectivity by measuring AM detection thresholds in the presence of a noise AM masker. These measurements also showed a bandpass characteristic with approximately constant filter bandwidth relative to the filter center frequency (constant Q value).

Grantham (1984) also reported a low-pass shape in his ILD modulation detection thresholds. Those data were obtained through measurements of the threshold of *discriminability* of interaurally antiphasic AM from homophasic AM imposed on interaurally uncorrelated, bandpass noise carriers. In contrast to the diotic TMTFs described earlier, the modulation depths required to discriminate the ILD modulation with low-frequency AM were quite high, around $m=0.15$ (−16 dB). In another study, Grantham and Bacon (1991) measured monaural and ILD modulation *detection* thresholds with broadband noise carriers and unmodulated reference intervals. Their monaural AM detection thresholds were very similar to those from Viemeister (1979) with thresholds of around −28 dB for low modulation frequencies. The ILD modulation detection thresholds were almost identical to the monaural thresholds, thereby showing 12 dB greater sensitivity to the modulation than reported by Grantham (1984). This increase in performance can be attributed to the difference in paradigm (AM detection versus discrimination). Since relatively small AM depths can be detected monaurally, characterization of binaural processing of modulated ILDs should only be done with the elimination of the monaural AM cues through an AM discrimination paradigm [as done by Grantham (1984)].

Grantham and Bacon (1991) also measured monaural and binaural frequency tuning in the envelope domain by measuring the TMTF in the presence of an AM masker. One diotic broadband-noise carrier, with a diotic tonal or narrow-band-noise amplitude modulator (the masker), was added to a second diotic broadband-noise carrier with an interaurally antiphasic sinusoidal amplitude modulator (the signal). They reported a bandpass tuning in the masked detection thresholds with the tonal modulator, but could not conclude whether that tuning was due to monaural or binaural processing. However, with a noise modulator masker, they did not see an effect of masker bandwidth and this led them to argue against a binaural modulation frequency tuning. Grantham (1984) described diotic AM as creating an "up-and-down flutter" with perceived changes in level or roughness, and antiphasic AM as creating a "side-to-side flutter" with a perception of motion or broadening between the ears. Assuming that the detection of the signal interval in their 1991 study was based on a comparison of the perceived motion or width of the two presentation intervals, a *diotic* masker, with no interaural fluctuations itself, would be perceived as motionless and narrow, and should have had little effect on the detectability of the modulated ILD signal. A *dichotic* masker, which does generate interaural fluctuations, would be better to measure the masked sensitivity to ILD modulations, along with a task of discriminating between interaurally antiphasic and homophasic AM ($AM_\pi - AM_0$), where the monaural cues have been made ambiguous. In this way, the results and any modulation frequency tuning could be attributed to purely binaural processing.

Stellmack *et al.* (2005) measured the sensitivity to ILD modulations using high-frequency (5 kHz) pure-tone and narrow-band noise carriers (30 and 300 Hz wide, interaurally correlated and uncorrelated), and an $AM_\pi - AM_0$ discrimination task. The thresholds measured with the pure-tone carrier were approximately constant at about −20 dB up to about $f_m=100$ Hz, where the sensitivity worsened with increasing f_m until the threshold could no longer be determined above $f_m=500$ Hz. There was a small increase in thresholds (up to 7 dB with the 30-Hz-wide carrier) when using the correlated noise carriers, relative to the pure-tone carrier data, particularly with low modulation rates ($f_m<20$ Hz). This increase was much smaller than the increase in thresholds seen with uncorrelated noise carriers, and was described as independent of the intrinsic carrier fluctuations. Therefore, the main focus of their paper was on the thresholds measured with uncorrelated narrow-band noise carriers.

Narrow-band Gaussian noises fluctuate randomly with envelope frequencies up to the bandwidth of the noise (see, e.g., Lawson and Uhlenbeck, 1950; Price, 1955). Monaurally, those inherent fluctuations can make it more difficult to detect an imposed AM, as compared to AM imposed on a pure-tone (i.e., flat envelope) carrier, especially when the AM frequency is less than the bandwidth of the noise carrier. This increase in threshold can be viewed as the result of masking of the signal AM by the intrinsic envelope fluctuations of the carrier. Binaurally, presenting interaurally uncor-

related narrow-band noises to each ear creates a dynamic ILD and the perception of a randomly moving or broad sound, depending on the bandwidth. The modulation spectrum of the ILD fluctuations is governed by the frequency content of the envelopes of the stimuli. The difference between the ILD modulation thresholds measured with uncorrelated and correlated noise carriers by Stellmack *et al.* (2005) also showed that the intrinsic ILD fluctuations from the uncorrelated carriers had the largest effect on thresholds for AM frequencies within the bandwidth of the noise. This suggests that there might be a frequency-selective mechanism in the processing of ILD fluctuations, similar to the monaural modulation AM processing from Dau *et al.* (1997a), but with broader frequency tuning.

The goal of the present study was to further investigate the modulation frequency tuning of the processing of ILD fluctuations. In the first experiment, detailed in Sec. III, the measurements of sensitivity to modulated ILDs from Stellmack *et al.* (2005) with narrow-band noise carriers were repeated with an additional carrier bandwidth (3 Hz, added to the 30- and 300-Hz-wide carriers) and a lower modulation frequency range (2–128 Hz instead of 4–600 Hz). A 3-Hz-wide Gaussian noise carrier has intrinsic modulations that can easily be followed (as loudness fluctuations monaurally, or as motion interaurally), where the 30- and 300-Hz-wide carriers are perceived with more roughness or width from the higher intrinsic modulation frequencies. The addition of the 3-Hz-wide carrier also enabled a comparison with all three carrier bandwidths (3, 31, and 314 Hz wide) used by Dau *et al.* (1997a). With the additional data from the present study, a different interpretation of the results than that of Stellmack *et al.* is proposed, which includes more emphasis on the threshold differences with diotic carriers.

Section IV details a second experiment for directly measuring the modulation frequency tuning for ILD fluctuations, using experimental design elements from Bacon and Grantham (1989), Houtgast (1989), Ewert and Dau (2000), and Ewert *et al.* (2002). In addition, simulations were made with an existing binaural computational model (from Breebaart *et al.*, 2001a), which was designed mainly for static interaural conditions, but includes a sliding integrator window (low-pass filter) to limit the temporal resolution. This enables it to predict some signal detection thresholds with dynamic interaural conditions (see Breebaart *et al.*, 2001c). The simulations should show whether an existing binaural model can predict similar thresholds to those of a human listener when used as an artificial observer.

II. GENERAL METHODS

Two psychoacoustic experiments were performed in order to investigate the sensitivity of the binaural system to modulated ILDs. In both experiments, the listener's task was to discriminate between a stimulus with interaurally antiphasic AM (AM_π; subscript indicating the interaural modulation phase) and a stimulus with homophasic (diotic) AM (AM_0). The AM frequency and depth was the same in all stimulus intervals of a three-interval, three-alternative forced-choice (3-AFC) trial, with only an interaural difference in modulation phase in the signal interval. The stimuli were defined as in Eq. (1) with carriers x_L and x_R (subscripts L and R for left and right ears, respectively):

$$\text{Left: } [1 + m\sin(2\pi f_m t + \phi_L)]x_L(t),$$

$$\text{Right: } [1 + m\sin(2\pi f_m t + \phi_R)]x_R(t), \tag{1}$$

where m is the modulation depth, f_m is the modulation frequency, and $\phi_{L/R}$ is the initial modulation phase for the respective ear's stimulus. The reference intervals were defined by Eq. (1) with $\phi_R = \phi_L$ (AM_0) and the signal interval was defined with $\phi_R = \phi_L + \pi$ (AM_π). The instantaneous ILD of a stimulus is defined as the ratio of the envelopes (as in Stellmack *et al.*, 2005):

$$\mathrm{ILD}(t) = 20\log_{10}\left(\frac{E_L(t)}{E_R(t)}\right). \tag{2}$$

A diotic sound does not create any ILDs itself. Therefore, the instantaneous ILD with an AM_π signal imposed on a diotic carrier is simply the ratio of the modulators:

$$\mathrm{ILD}(t) = 20\log_{10}\left(\frac{1 + m\sin(2\pi f_m t + \phi_L)}{1 - m\sin(2\pi f_m t + \phi_L)}\right) \tag{3}$$

and the maximum ILD is a function of the modulation depth only:

$$\mathrm{ILD}_{\max} = 20\log_{10}\left(\frac{1+m}{1-m}\right). \tag{4}$$

Interaurally uncorrelated noises produce stochastic ILD fluctuations, which add linearly (on a decibel scale) to the deterministic signal ILD modulation. These random ILD fluctuations will change the distribution of ILDs and the maximum ILD of the stimulus.

The two experiments differed in specifics of the stimuli (e.g., modulation phase and carrier), which will be presented in Secs. III and IV, but the general methods were the same.

A. Test subjects

Four test subjects were used for all experiments. They were not paid directly for their participation, but were all involved at the research center, and included both authors of this paper. All had pure-tone audiometric thresholds of 15 dB HL or better for octave frequencies between 250 Hz and 8 kHz. They were all experienced in psychoacoustic measurements and particularly in AM detection experiments. That experience was mostly with monaural or diotic stimuli, so their experience with the detection of interaural fluctuations was limited. All listeners were encouraged to listen to example stimuli and performed a limited set of training runs of approximately 1 h duration.

B. Equipment

All signals were generated and presented using the AFC-Toolbox for MATLAB (Math-Works), developed at the University of Oldenburg, Germany and the Technical University of Denmark, at a sample rate of 44.1 kHz through a sound card (RME DIGI 96/8 PAD) and headphones (Sennheiser

HD-580). The test subjects sat in a sound insulated booth with a computer monitor, which displayed instructions and visual feedback, and a keyboard for response input.

C. Procedure

A 3-AFC paradigm was used with an adaptive one-up, two-down tracking rule, which should converge at the 70.7% correct point on the psychometric function (Levitt, 1971). In a given track, the modulation frequency of the AM signal was fixed and the modulation depth was varied to find the AM depth required for identification of the signal. During each trial, a computer monitor displayed a window with three buttons, representing the three stimuli. Each button was highlighted when the corresponding interval was played. The signal interval was randomly selected with equal probability of occurrence from the three presentation intervals. The test subject responded via the computer keyboard and received immediate feedback on whether the response was correct or incorrect. All tracks were assembled in one long experiment that the test subject could start and stop at will after the completion of any track. The typical duration of a session was about 30 min, but could be longer or shorter depending on the circumstances.

Each track started with a modulation depth of −2 dB ($20\log_{10} m$). The step size started at 4 dB, and was halved after every second reversal until the final step size of 1 dB was reached after the fourth reversal. The track continued for six further reversals with this step size, and the threshold was determined as the mean of those last six reversals. Each test subject completed four repetitions for each modulation frequency and set of experimental parameters, and the results shown are the mean and standard deviation of all test subjects and repetitions. If the test subject could not identify the correct interval with the maximum modulation depth ($m=0$ dB) twice in a row, the track was skipped and the experiment continued to the next track.

D. Common stimulus parameters

All stimuli were centered at 5 kHz, so that all frequency components would lie well above the range of frequencies in which interaural timing differences in the fine structure of the carriers would affect the lateralization of the stimuli (e.g., Klumpp and Eady, 1956; Mills, 1960). The stimuli were gated simultaneously in the two ears, and presented at a level of 65 dB SPL, with 300 ms of silence between intervals. In tracks where noise carriers were used, a new noise sample was generated for each presentation.

III. EXPERIMENT I: MODULATION DISCRIMINATION WITH NARROW-BAND NOISE CARRIERS

The first experiment was designed to measure the sensitivity of binaural processing to modulated ILDs, using an experimental design based on the diotic AM detection measurements from Dau *et al.* (1997a). Parts of this experiment are a repetition of similar experiments performed by Stellmack *et al.* (2005).

A. Specific stimulus details

ILD modulations were created by applying an interaurally antiphasic sinusoidal AM to pure-tone and narrow-band noise (3-, 30-, and 300-Hz-wide) carriers, centered at 5 kHz. In order to eliminate monaural modulation cues, an $AM_\pi - AM_0$ discrimination paradigm was used. The noise carriers were generated by creating a 1-s-long independent Gaussian noise sample for each interval in the time domain and setting the frequency components outside of the passband to zero in the spectral domain. Measurements were made with interaurally correlated (symbol N_0) and uncorrelated (N_u) noise carriers. Sinusoidal AM was applied to the carrier as given in Eq. (1) with $\phi_L=\phi_R=0$ (AM_0) in the reference intervals and $\phi_L=0$ and $\phi_R=\pi$ (AM_π) in the signal interval. With this choice of AM phase parameters, the change in modulation phase in the right ear could have been used as a monaural cue for signal detection. Therefore, a control experiment was performed to measure the modulation depth required for discrimination of monaural AM phase change (referred to as AM_m disc) using a pure-tone carrier and the right ear only. For these tracks, the signal interval had an initial modulation phase of π (negative-going zero crossing) and the reference intervals started with a modulation phase of zero (positive-going zero crossing).

Thresholds were measured with AM frequencies (f_m) in octave steps from 2 to 32 Hz and 128 Hz. At the highest modulation frequency used (128 Hz), an interaural modulation phase difference of π is equivalent to an ITD of ± 3.9 ms. Since this is well above the ecologically relevant range of ITDs [approximately 650 μs for humans, Feddersen *et al.* (1957)], and because the test subjects reported informally hearing a diffuse sound image, and not a static lateralization of the sound, it is assumed that listeners did not use the static envelope ITD to localize the sound source. Therefore, the experiments will be discussed in terms of dynamic ILDs. Tracks with $f_m \geqslant 8$ Hz had a stimulus duration of 500 ms, including 50 ms $\cos^2$ onset and offset ramps, while tracks with $f_m=2$ or 4 Hz had a duration of 1000 ms in order to reduce the interference of the windowing function on the desired envelope frequency components [the stimulus windows of Stellmack *et al.* (2005) were 1 s long with 150 ms ramps]. The intervals were separated by 300 ms of silence, where the intervals of Stellmack *et al.* (2005) were only demarcated by the ramps with no additional separating silence. Previous measurements from Dau (1996) showed that listeners could not reliably discriminate (defined there as $P_c > 33\%$) between monaural AM phase with full modulation ($m=1$) for $f_m > 12$ Hz, using a 5 kHz pure-tone carrier. Therefore, the AM_m phase discrimination threshold was only measured in the present study with $f_m \leqslant 8$ Hz. More recent results from Sheft and Yost (2007) showed that some listeners could only discriminate modulation starting phase with broadband noise carriers up to about $f_m=12.5$ Hz, while others were still able to perform the task up to around 50 Hz.

Two additional control measurements were made with an AM *detection* paradigm, where only the signal interval in a three-interval trial had an applied AM and the two reference intervals were unmodulated. Monaural AM_m and inter-

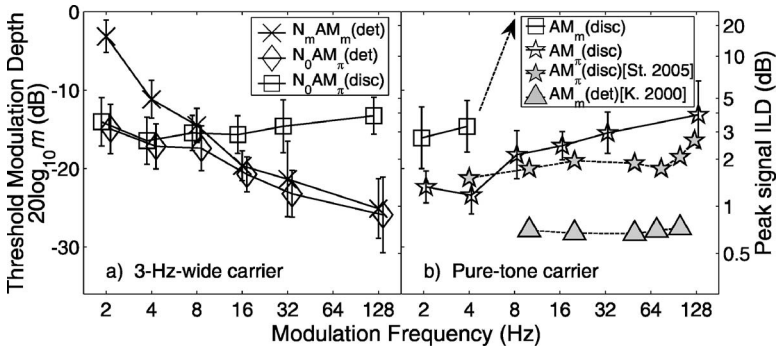

FIG. 1. Amplitude modulation detection and discrimination thresholds in decibels for various monaural and binaural conditions. The left ordinate labels (modulation depth) apply to all curves in both panels, and the right ordinate labels (peak ILD) applies to the binaural (AM_π) thresholds in both panels. (a) Thresholds measured with a 3-Hz-wide noise carrier. Monaural AM detection (N_mAM_m, X's), ILD modulation detection (N_0AM_π, diamonds), and ILD modulation discrimination (N_0AM_π–N_0AM_0, squares). (b) Thresholds measured with a pure-tone carrier for monaural AM phase discrimination (AM_m, squares) and ILD modulation discrimination (AM_π–AM_0, open stars). The data shown with shaded symbols are ILD modulation discrimination [shaded stars, adapted from Stellmack *et al.* (2005), Fig. 3], and monaural AM detection thresholds [shaded triangles, adapted from Kohlrausch *et al.*, (2000), Fig. 2, 5 kHz carrier]. Note that the data in (a) are offset around the AM frequency for visual clarity of the error bars.

aural AM_π detection thresholds were measured with a 3-Hz-wide carrier in order to demonstrate the importance of eliminating monaural cues in the binaural experiment. Diotic noise carriers were used for this AM_π detection measurement.

B. Results

The results from the four test subjects were similar in shape and value, so the plots shown in Figs. 1–3 display the mean and standard deviation over all test subjects and runs. The modulation depths required for detection or discrimination of monaural AM and modulated ILDs are plotted in decibels ($20\log_{10}m$) as a function of the signal modulation frequency for the pure-tone [Fig. 1(b)] and narrow-band noise (Fig. 2) carriers. Note that the ordinates are shown with larger modulation depths and therefore poorer sensitivity toward the top of the axis. In Figs. 1 and 2, there is also a second ordinate on the right of each plot showing the peak *signal* ILD in decibels, which applies to all of the AM_π data. This peak ILD was calculated from the modulation depth at threshold according to Eq. (4). As discussed earlier, the actual peak ILD seen with the uncorrelated noise carriers varied around this value. In the following, the results are presented in terms of the modulation depth at threshold in decibels, unless otherwise noted. The data points are offset slightly from the AM frequency in Figs. 1(a), 1(b), 2(b), and 2(d) for visual clarity of the error bars. A two-way analysis of variance (ANOVA) with repeated measures was used to

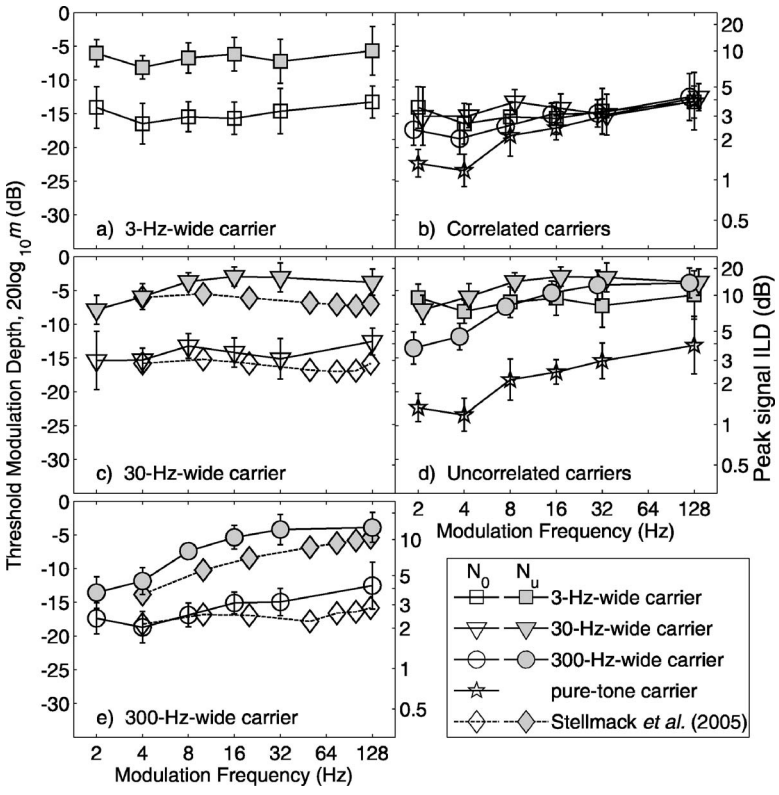

FIG. 2. ILD modulation discrimination thresholds measured with narrow-band noise carriers. The left and right ordinate labels apply to all curves in all panels, with the AM depth (left ordinate) converted to peak ILD in the right ordinate according to Eq. (4). In all panels, the interaural correlation of the carriers is indicated by the shading. Open symbols indicate correlated carriers and shaded symbols indicate uncorrelated carriers. The symbols indicate the carrier bandwidth: Squares for 3 Hz wide, triangles for 30 Hz wide, and circles for 300 Hz wide. The pure-tone thresholds from Fig. 1(b) are replotted with open star symbols in (b) and (d). External data from Stellmack *et al.* (2005) are indicated by diamonds for both 30- and 300-Hz-wide data. The thresholds for the 3-, 30-, and 300-Hz-wide carriers are grouped by band-width in (a), (c), and (e), respectively. The data measured with correlated carriers are replotted in (b), and with uncorrelated carriers in (d). Note that the data in (b) and (d) are offset around the AM frequency for visual clarity of the error bars.

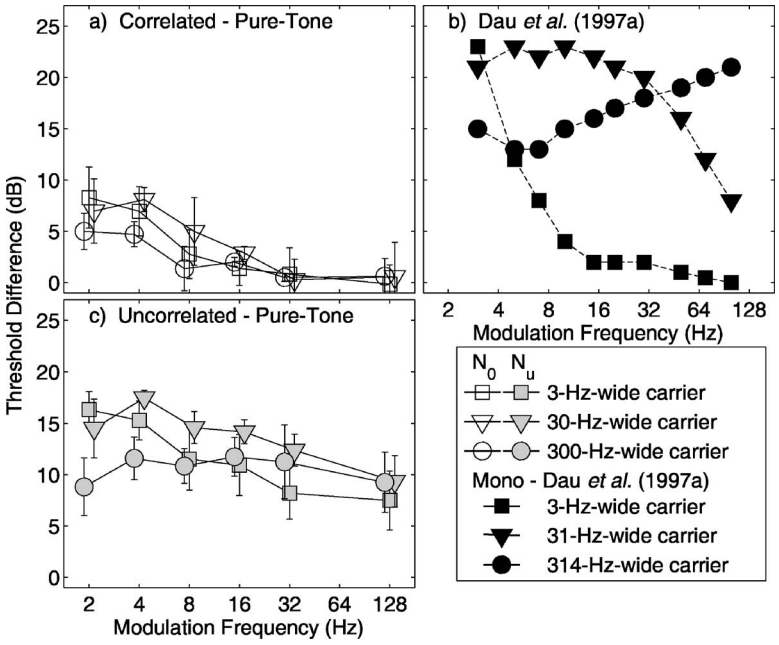

FIG. 3. The difference in discrimination thresholds measured with narrow-band noise carriers and pure-tone carriers [AM_π, stars from Fig. 1(b)]. (a) The difference for thresholds measured with correlated noise carriers [N_0AM_π, from Fig. 2(b)]. (b) The difference between monaural AM detection with narrow-band noise and pure-tone carriers [data adapted from Dau *et al.* (1997a)]. (c) The difference for thresholds measured with uncorrelated noise carriers [N_uAM_π, from Fig. 2(d)]. The error bars in (a) and (c) show the standard deviation of the mean threshold difference across listeners.

compare the data curves with a threshold of $p<0.05$ required for significance.

Figure 1(a) shows the thresholds obtained for the monaural N_mAM_m detection (X's), N_0AM_π detection (diamonds), and N_0AM_π discrimination (squares) with a 3-Hz-wide carrier. The monaural curve shows the AM frequency selectivity described by Dau *et al.* (1997a) in that the highest thresholds (−3 dB at $f_m=2$ Hz) are at a frequency within the bandwidth of the noise (i.e., $f_m<3$ Hz), and the thresholds for AM detection decrease with increasing f_m (down to −25 dB at $f_m=128$ Hz). Discrimination of AM_π from AM_0 with the 3-Hz-wide carrier is approximately constant with thresholds between −16 and −13 dB for all measured f_m. This shows that while the low-frequency modulations ($f_m<8$ Hz) are obscured by the fluctuations of the carrier in each ear (up-and-down), the ILD modulation (side-to-side) is still easily detectable with harmonic ILD oscillations with amplitudes of 2.8–4 dB [see the right ordinate in Fig. 1(a)]. The N_0AM_π *detection* threshold demonstrates how the auditory system switches from monaural to binaural cues, depending on cue salience. For low AM frequencies ($f_m<8$ Hz), where the monaural cues are obscured by the envelope fluctuations of the 3-Hz-wide carrier, there is no significant difference between the N_0AM_π detection and discrimination thresholds. For $f_m>8$ Hz, where the carrier fluctuations have a relatively small influence on the monaural detectability, the N_0AM_π detection and N_mAM_m thresholds have no significant differences.

In the control experiment to ensure that the test subjects could not perform the binaural discrimination tasks based solely on a monaural AM phase cue, thresholds for monaural AM phase discrimination could only be measured for $f_m\leqslant4$ Hz [squares in Fig. 1(b)]. At $f_m=8$ Hz, the discrimination task could not reliably be performed by any of the test subjects, even at full modulation depth ($m=0$ dB) and the arrow indicates that no threshold was measurable. These data correspond well to those from Dau (1996) and with some of the listeners from Sheft and Yost (2007), but cannot help to explain how some listeners in the latter study were still able to discriminate modulation starting phase at rates up to about 50 Hz. The thresholds for monaural modulation phase discrimination in the present study for the 2- and 4-Hz AM signals showed which of the results from the other measurements presented in this experiment (Figs. 1 and 2) could have been influenced by a monaural modulation phase cue. Those measurements were repeated informally with only the right ear's signal to verify that the tasks could not be performed monaurally at the measured threshold levels, and none of the listeners tested were able to do so.

The ILD modulation discrimination threshold curve measured with a pure-tone carrier [open stars in Fig. 1(b)] shows an overall low-pass tendency with thresholds around −23 dB (1.2 dB peak ILD) for $f_m=2$ and 4 Hz and increasing to about −13 dB (4 dB peak ILD) for $f_m=128$ Hz. Stellmack *et al.* (2005) reported a flatter threshold shape with thresholds around −20 dB (1.7 dB peak ILD) from $f_m=4$ to almost 100 Hz [shaded stars in Fig. 1(b)], above which the threshold increased. The monaural TMTF with a pure-tone carrier [shaded triangles in Fig. 1(b)], reported by Kohlrausch *et al.* (2000), shows that the auditory system is much more sensitive to envelope fluctuations (thresholds around −28 dB up to $f_m>100$ Hz) than to the ILD fluctuations caused by an interaural modulation phase inversion.

The data measured with narrow-band noise carriers (3, 30, and 300 Hz wide) are plotted twice in Fig. 2. The left panels show the data grouped by carrier bandwidth, and the right panels by carrier interaural correlation. In all panels, the squares represent data for the 3-Hz-wide carrier, the triangles for the 30-Hz-wide carrier, and the circles for the 300-Hz-wide carrier data. Open symbols indicate interaurally correlated carriers (N_0), and shaded symbols are for interaurally uncorrelated carriers (N_u). In addition, the corresponding data from Stellmack *et al.* (2005) for the 30- and 300-Hz-wide carriers are shown in Figs. 2(c) and 2(e), re-

spectively, with diamonds and dashed lines, and the pure-tone (AM_π) thresholds are replotted from Fig. 1(b) in Figs. 2(b) and 2(d) with stars. The 30- and 300-Hz-wide carrier data show a good agreement with the data from Stellmack *et al.* (2005), with only small differences that could be the result of the procedural differences discussed earlier (e.g., stimulus length, windowing).

By grouping the thresholds by carrier bandwidth [Figs. 2(a), 2(c), and 2(e)], a strong effect of the interaural carrier correlation can be seen. The N_uAM_π thresholds are much higher than the N_0AM_π thresholds. With the 3-Hz-wide carriers [Fig. 2(a)], the N_uAM_π thresholds show an almost constant offset of about 8 dB from the N_0AM_π curve. The differences between the thresholds measured with wider bandwidth carriers [Figs. 2(c) and 2(e)] increase with increasing modulation frequency from 7 to 12 dB with the 30-Hz-wide carriers and from 4 to 11 dB with the 300-Hz-wide carriers.

A two-way ANOVA with repeated measures was calculated on the N_0AM_π data [Fig. 2(b)] with carrier bandwidth and modulation frequency as factors. The analysis showed no significant effect of bandwidth ($p=0.13$), but a significant effect of modulation frequency ($p<0.05$) and a significant interaction between bandwidth and modulation frequency ($p<0.05$). Adding the pure-tone AM_π data to the analysis as an additional bandwidth yielded a significant effect of bandwidth ($p<0.01$). An ANOVA on the data measured with interaurally uncorrelated noise carriers [N_uAM_π; Fig. 2(d)] showed a significant effect of bandwidth, even without the pure-tone carrier data, and of modulation frequency, as well as a significant interaction between the factors ($p<0.01$ for both factors and the interaction).

C. Discussion

The modulation depths required to discriminate AM_π from AM_0 imposed on diotic noise carriers were significantly larger than those required with pure-tone carriers, particularly with low modulation rates ($f_m<16$ Hz). The discrimination thresholds measured with interaurally uncorrelated noise carriers were even higher than those measured with the correlated noise carriers. The increase in thresholds when using noise carriers instead of pure-tone carriers can be considered as masking of the modulation signal by the intrinsic envelope fluctuations of the noise carriers themselves, since a pure-tone carrier does not have these random fluctuations. Figure 3(a) shows the *difference* in thresholds measured with correlated noise carriers and the pure-tone carrier. Note that this reflects an increase in the threshold of ILD modulation discrimination as the result of *diotic* fluctuations of the noise carriers, which do not create ILD fluctuations themselves. Looking at the differences between the N_uAM_π thresholds and the pure-tone carrier thresholds [plotted in Fig. 3(c)], the 3- and 30-Hz-wide carriers show the greatest difference for f_m below the bandwidth of the carrier, beyond which the difference decreases monotonically toward an asymptotic value of about 7–9 dB. The threshold difference with the 300-Hz-wide carrier increases from about 8 dB with $f_m=2$ Hz to 12 dB with $f_m=16$ Hz, and then decreases again to 9 dB at $f_m=128$ Hz.

Stellmack *et al.* (2005) also observed an increase in $AM_\pi-AM_0$ discrimination thresholds with the diotic 30- and 300-Hz-wide carriers, but did not report an effect of the carrier-envelope frequency content on the shape of the increase, and therefore focused on the difference between the thresholds with uncorrelated and correlated noise carriers. However, there is a significant interaction between the diotic noise bandwidths and the modulation frequency in the present study, which can be seen in the threshold differences for $f_m<16$ Hz [see Fig. 2(b)]. These differences could suggest a dependence on the carrier envelope spectrum, but this needs to be investigated in further studies.

A control experiment was performed to investigate the difference between the thresholds measured in the AM_π discrimination experiments with the pure-tone and the diotic narrow-band noise carriers. The ILD modulation discrimination threshold was measured with a 30-Hz-wide "low-noise noise"[1] (Pumplin, 1985) carrier. Measuring with a low-noise noise carrier tests the hypothesis that the difference in thresholds between AM_π discrimination with a pure-tone carrier and with a narrow-band noise carrier is caused by the envelope fluctuations of the noise and not by the noise's broader bandwidth per se. A pairwise t-test of the results showed no significant difference between thresholds measured with the low-noise noise carrier and the pure-tone carrier (thresholds at −19.0 and −17.8 dB, respectively, $p=0.43$), but there was a significant difference between the low-noise noise and the Gaussian noise thresholds (−19.0 and −14.7 dB, respectively, $p<0.05$). This suggests that it is the fluctuations in level of the Gaussian noise carriers that impede the discrimination of AM_π from AM_0.

The diotic Gaussian noise carrier and the pure-tone carrier do not create any ILDs themselves. Therefore, the binaural sensitivity to the AM_π signal is only limited by the internal variability of the auditory system, or by "internal noise." Since there is a significant increase in thresholds when using diotic Gaussian noise carriers, this suggests that the internal noise increases when the envelopes of the carriers fluctuate, or that the encoding of fluctuating envelopes is noisier than the encoding of steady envelopes. The range of modulation frequencies over which there is an increased threshold with the correlated noise carriers, and the differences between the thresholds from the three carrier bandwidths should provide some insight into how the interaural processing differences can be modeled.

Interaurally uncorrelated Gaussian noise carriers cause large stochastic fluctuations in ILD. This "external" ILD variability is in addition to the internal noise described earlier. Therefore, it is not surprising that the ILD modulation discrimination thresholds are higher with N_uAM_π than with N_0AM_π. It is unknown how the effects of the external and internal variances with Gaussian noise carriers combine in the auditory system. Therefore, the data obtained with *pure-tone* carriers were used as the reference threshold in the present study. This is in contrast to Stellmack *et al.* (2005), who compared the thresholds with uncorrelated and correlated noise carriers, leaving out the extra effect of the diotic

level fluctuations on the measurements. The thresholds measured with the uncorrelated carriers are up to 18 dB higher than those measured with the pure-tone carrier, particularly at AM frequencies below the bandwidth of the carrier. By comparing with the pure-tone carrier thresholds instead of with the diotic noise carriers, the shapes of the threshold difference curves with uncorrelated noise carriers from Fig. 3(c) have similar aspects to those from monaural experiments [Fig. 3(b); adapted from Dau *et al.* (1997a)], but also large differences. The monaural curves [Fig. 3(b)] with 3 and 31-Hz-wide carriers drop off quickly toward zero with f_m greater than the carrier bandwidth, indicating relatively sharp modulation frequency tuning. The binaural curves [Fig. 3(c)] also roll off with f_m greater than the carrier bandwidth, but do not roll off as quickly as the monaural curves, and seem to reach a plateau at about 8 dB, even with f_m much greater than the carrier bandwidth. This indicates much broader tuning in the binaural domain than in the monaural domain, as also suggested by Stellmack *et al.* (2005). In contrast, Grantham and Bacon (1991) argued against a bandpass ILD modulation tuning after measuring detection thresholds with a 16-Hz AM_π signal in the presence of a diotic noise AM masker. At that AM frequency, the data in the present study show no significant differences between the AM_π discrimination thresholds with correlated noise or pure-tone carriers, even though the 30- and 300-Hz-wide carriers have envelope frequency components around 16 Hz. This suggests that the diotic masker in their study probably did not have a significant effect on the AM_π detection threshold. This analysis suggests that the data from Grantham and Bacon (1991) is equivocal on the presence or absence of bandpass ILD modulation tuning.

The qualitative similarities between the binaural and monaural masking curves suggest that an element could be introduced in a binaural model that is similar to the monaural modulation filterbank from Dau *et al.* (1997a). However, it appears that the tuning of the ILD modulation filters in such a model must be broader than those of the monaural filterbank.

IV. EXPERIMENT II: MASKED MODULATION DETECTION

The results of the first experiment suggested that there may be modulation frequency selectivity in the processing of ILD fluctuations. Therefore, further experiments were performed to directly measure the shape of this tuning. These experiments were based on similar experiments performed with diotic signals by Ewert *et al.* (2002), where a sinusoidal signal AM was masked by a narrow-band noise modulator applied to a common pure-tone carrier.

A. Specific stimulus details

An interaurally uncorrelated, bandpass Gaussian-noise masker modulation was applied to the envelope of pure-tone carriers in a discrimination task, according to

$$x_L(t) = a\,\sin(2\pi f_c t)[1 + N_L(t)][1 + m\,\sin(2\pi f_m t + \phi_L)]\ ,$$

$$x_R(t) = a\,\sin(2\pi f_c t)[1 + N_R(t)][1 + m\,\sin(2\pi f_m t + \phi_R)]\ , \tag{5}$$

where a controls the presentation level, f_c is the carrier frequency (in this case, 5 kHz), the subscripts L and R indicate left or right ear, and $N_{L/R}$ is the masking noise modulator (power set in this study to −10 dB re 1), spectrally centered at f_N, for the respective ears. The signal modulation was applied with AM frequency f_m, modulation depth m, and starting phases ϕ_L and ϕ_R. When two amplitude modulators are to be applied to a carrier (e.g., a masker, N, and a signal modulator, S), they can be added together and applied as a common modulator $(1+S+N)$ or applied in series as separate modulators $(1+S)(1+N)$. The additive approach can result in overmodulation if either the signal or the masker has a large negative amplitude (i.e., $S+N<-1$). The multiplicative approach, used in this study, avoids overmodulation as long as $S>-1$ and $N>-1$, which allows for signal modulation depths (m) close to 0 dB (see also Houtgast, 1989). However, by multiplying the two modulators, additional spectral sidebands are created, which can complicate analysis of the data, as discussed in Sec. IV C.

The design of the stimuli was based on Ewert *et al.* (2002). Each stimulus had an overall duration of 600 ms, windowed with 50 ms $\cos^2$ onset and offset ramps. The AM signal was applied to the middle 500 ms of the carrier, gated with 50 ms $\cos^2$ onset and offset ramps, leaving 400 ms with the desired signal AM depth. Measurements were made with f_m=4, 8, and 32 Hz. In order to avoid monaural cues, the experiment was designed as a discrimination task, so all three intervals in a trial (signal and two references) had an applied signal modulation with the same modulation depth in each interval. The start phase of the signal modulation in the left ear ϕ_L was chosen randomly for each trial interval over the range $[0, 2\pi]$ with a uniform probability distribution. In the two reference intervals, the modulation start phase in the right ear was set equal to ϕ_L (AM_0), while in the signal interval, ϕ_R was set equal to $\phi_L+\pi$ (AM_π). With the randomized modulation phase and equal modulation depth on all intervals in a trial, successful discrimination could only be performed by combining information from the two ears, not based on one ear's analysis alone.

The masker modulations had a fixed bandwidth of 1.4, 2.8, and 11.1 Hz for the f_m=4, 8, and 32 Hz, respectively, corresponding to one half-octave centered at f_m. The masker center frequencies, f_N, were at octave steps from f_m over a range from −4 to +4 octaves, but with the additional limitation that f_N could not be below 2 Hz or above 128 Hz. This hard frequency limit was put in place because the envelope of the window function itself could interfere with detection below 2 Hz, and the modulation sidebands could be resolvable above 128 Hz. In Ewert *et al.* (2002), the masker modulation was placed over a range from −2 to +2 octaves with a 2/3 octave step size. The larger range and step size were chosen here in expectation of broader tuning after the results from Experiment 1 (see Sec. III). A new masker modulator $(N_{L/R})$ was created for each presentation interval by generat-

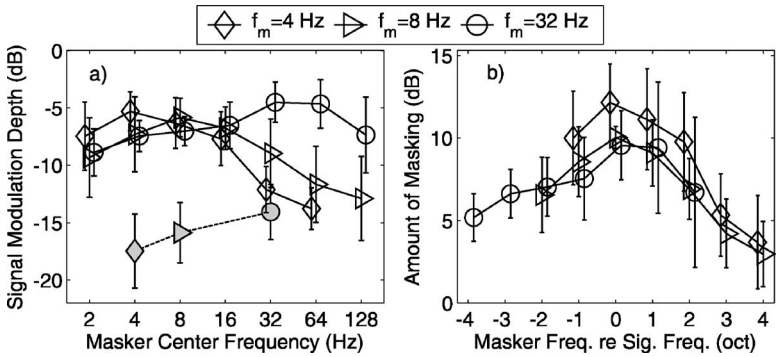

FIG. 4. (a) Modulation depths required for discrimination of interaurally antiphasic AM from homophasic AM imposed on a pure-tone carrier in unmasked (shaded symbols, dashed line) and masked (open symbols, solid lines) conditions. Measurements were made with a fixed signal AM frequency of 4 Hz (diamonds), 8 Hz (triangles), and 32 Hz (squares) with interaurally uncorrelated narrow-band noise maskers with a fixed power and bandwidth for a range of masker center frequencies. The data points for each curve are offset to more clearly show the error bars. (b) The same data from the left panel, but normalized for the unmasked threshold and signal frequency. The error bars in (b) show the standard deviation of the mean threshold difference across listeners.

ing a 10 s Gaussian white noise in the time domain, setting all frequency components outside the passband to zero, and then scaling the variance to 0.1 (−10 dB re 1). The resulting noise was then added to a dc component $(1+N_{L/R})$ and applied to the carrier as in Eq. (5). At this masker level, there was a small probability (less than 0.08% of samples, or less than 0.5 ms per presentation, on average) of overmodulation (i.e., $1+N_{L/R}<0$). This small occurrence was assumed to not have a significant effect on the results.

B. Results

Figure 4(a) shows the mean and standard deviation of the masked threshold patterns measured with a pure-tone carrier. The signal modulation depth in decibels $(20 \log m)$ is plotted as a function of the masker center frequency, with the signal modulation frequency as the parameter. In addition, the modulation depth required for discrimination without a masker present is plotted as a function of the signal modulation frequency (dashed line, shaded symbols). Note that the three masked curves and their respective unmasked points have been offset slightly around the exact frequencies so that the error bars are more visible. In Fig. 4(b), the same curves are replotted as an amount of masking, defined as the difference between the masked and unmasked thresholds for each signal modulation frequency at each masker center frequency, normalized by the signal modulation frequency in octaves. The error bars in Fig. 4(b) show the standard deviation of the mean threshold difference across listeners. In both panels, the symbols (diamond, triangle, and circle) represent the 4, 8, and 32 Hz signal modulation frequencies, respectively.

The masking patterns [Fig. 4(b)] for the three signal frequencies are very similar in shape and amount of masking. All three curves show the highest amount of masking (approximately 10 dB) for the on-frequency condition and a monotonic decrease in masking as the spectral distance between signal and masker center frequency increases. The decrease in masking is greater when the masker center frequency is above the signal frequency than with lower masker frequencies. When the masker is 4 oct. above the signal, there is only about 3 dB of masking, but there is still about 6 dB of masking with the masker 4 oct. below the signal.

C. Discussion

The masking patterns obtained in the $AM_{\pi}-AM_0$ discrimination task with a narrow-band noise modulator masker imposed in series with a sinusoidal signal modulator on a pure-tone carrier showed consistency in shape and amount of masking for the three measured signal AM frequencies. The mean values of the three masking curves at each relative masker frequency are replotted in Fig. 5 (circles) along with a typical masking curve (squares) from the monaural masked AM detection experiments of Ewert *et al.* (2002) (adapted from their Fig. 2; 5.5 kHz carrier, 64 Hz AM signal, $Q=1.25$). The two curves show a maximum amount of masking when the masker is centered at the signal frequency, although the monaural curve shows a clearly higher masking value (about 17 dB) than the binaural curve. For masker frequencies above the signal frequency, the monaural curve rolls off more rapidly than the binaural curve, so that the monaural curve already shows less masking than the binaural curve for maskers centered 1 oct. above the signal modula-

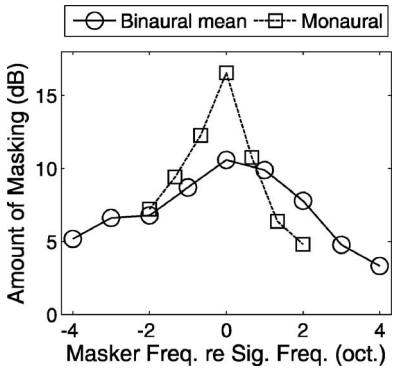

FIG. 5. The mean of the three masking curves from panel Fig. 4(b) (circles, solid line) is shown with a typical monaural AM masking curve [dashed line, squares, adapted from Ewert *et al.* (2002), from their Fig. 2, 5.5 kHz carrier, 64 Hz AM signal].

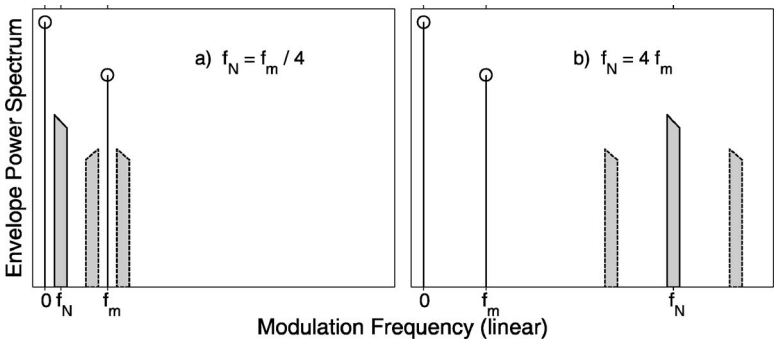

FIG. 6. Theoretical envelope power spectra resulting from the application of a band-pass noise masker modulator and a tonal signal modulator. The dc component ($f=0$) and tonal component ($f=f_m$) are plotted with circles and the spectrum of the noise masker is shown with a bar with a solid line centered at f_N. The bars shown with the dashed lines show the result of applying the masker and signal modulators in series with the multiplicative approach described in the text. The left panel shows a case where $f_N<f_m$ and the right panel shows a case where $f_N>f_m$. Note that only the dc component and positive frequencies are shown.

tion frequency. This is consistent with the idea that monaural envelope processing has a sharper tuning than binaural processing of dynamic ILDs.

Multiplication of the signal and masker modulators creates additional sidebands in the stimulus through spectral convolution. This is represented in a sketch of the envelope spectra of two idealized stimuli in Fig. 6. A stimulus with only an applied noise masker AM would show an envelope spectrum with a dc component ($f=0$) and a band of noise centered at f_N. Multiplication of the masking modulator with the signal modulator (tonal component at f_m) results in the two sidebands shown with dashed lines in Fig. 6 centered at $f_m \pm f_N$ (note that only positive frequencies are shown in the sketch). In an AM *detection* experiment, as in the monaural experiments from Houtgast (1989) and Ewert *et al.* (2002), where the listener's task is to distinguish between presentation intervals with only a masker modulator and a target interval with masker and signal modulators, the sidebands are only present in the target interval. Therefore, they can serve to enhance the detectability of the signal AM. However, with an AM *discrimination* experiment, like the one here, all stimuli have the same modulation and the same sidebands. In this case, the sidebands do not provide any cues for signal detection, and may actually hamper signal detection.

The amplitude of the sidebands is determined by the amplitudes of the masker and signal AM components. In this experiment, with a fixed masker energy, the sidebands' energy scales with the signal energy at a fixed ratio (−10 dB). The effect of these sidebands should be considered when designing a model to account for the measured masking patterns. For example, a model could be designed with a symmetric bandpass modulation filter centered at the signal's modulation frequency, and a certain signal-to-noise ratio (SNR) required after the filter for detection of the signal modulation. With this model, the sidebands' energy would create a noise floor at a SNR that depends on the signal and masker frequencies. When $f_N \ll f_m$, the sidebands will be very close, spectrally, to the signal [Fig. 6(a)], and the sidebands' energy will be passed through the filter with little attenuation, creating a relatively high noise floor. As f_N increases, the sidebands move away from the signal in frequency, becoming more attenuated by the filter and reducing the noise floor. The sidebands are only centered at frequencies larger than f_m if $f_N > 2f_m$. The result could be an asymmetric masking pattern, even though the filter was assumed to be symmetric around the signal modulation frequency.

V. IMPLICATIONS FOR BINAURAL MODELS

The above-presented experimental data suggest that a binaural model should include an array of ILD modulation bandpass filters to simulate human performance. Some preliminary simulations were made using the binaural model from Breebaart *et al.* (2001a) as an artificial observer in the experiments described in Sec. IV. These simulations were performed with the original model, which uses a sliding integrator (low-pass filter) to limit its temporal resolution. The Breebaart model was designed for static binaural conditions, such as for predicting binaural masking level differences (BMLD), and is quite successful at predicting human performance under many experimental conditions (see also Breebaart *et al.*, 2001b,c, for more details). Breebaart *et al.* (2001c) focused on temporal parameters, including a simulation based on an experiment from Grantham (1984), where the listener's task was to discriminate between interaurally antiphasic and homophasic AM imposed on uncorrelated broadband noise carriers. Grantham's data showed a large variance between test subjects, but Breebaart's model was able to capture the general trend of the results. Since their model was able to simulate experimental results similar to those described earlier in Sec. III, it was chosen as a basis for testing with the new experiments and for possible future development.

The model starts with two parallel peripheral processing stages (one for each ear, see schematic in Fig. 7), based on the monaural processing model from Dau *et al.* (1996), which did not include a modulation filterbank. The first stages are an outer- and middle-ear transfer function, basilar-membrane filtering, consisting of an array of gammatone filters, inner hair cell transduction, modeled with half-wave rectification and a low-pass filter with a cutoff frequency of 770 Hz, and finally a series of five adaptation loops, which enable the simulation of forward masking. The output from each pair (right/left) of peripheral channels is then passed to an array of excitation-inhibition (EI) elements, which calculate the difference in the corresponding channels for a range of characteristic interaural gains and delays. This is similar in concept to the equalization-cancellation (EC) model from Durlach (1963), which finds the optimal gain and delay before calculating the channel difference. The EI concept is based on neurons that receive excitatory input from the ipsilateral side and inhibitory input from the contralateral side, effectively calculating a difference between the two auditory signals. The output from each EI element is then smoothed

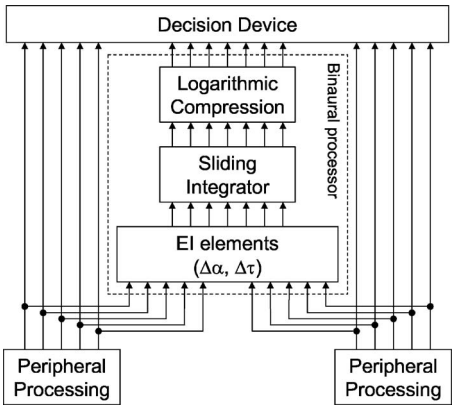

FIG. 7. Schematic of the binaural model from Breebaart *et al.* (2001a). The model consists of two parallel monaural peripheral channels, including a gammatone filterbank, a half-wave rectification and low-pass filter inner hair cell model, and adaptation loops. The two monaural signals are combined in the binaural processor through an array of excitation-inhibition (EI) elements, which calculate the difference of the two signals for a range of applied interaural gains and delays. The resulting signals are smoothed with a sliding integrator and compressed with a logarithmic compression. Finally, an optimal detector tries to find a signal based on all monaural and binaural inputs.

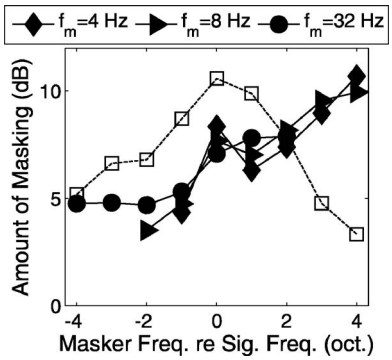

FIG. 8. Masked tuning curves predicted by the model from Breebaart *et al.* (2001a) when used as an artificial listener in the experiments described in Sec. IV. The mean tuning curves from the human listeners are shown with the dashed line and open squares [mean of the three curves from Fig. 4(b)]. The simulation was made with f_m=4 Hz (diamonds), 8 Hz (triangles), and 32 Hz (circles).

with a sliding integrator, consisting of a symmetric double-sided exponential window with time constants of 30 ms. This sliding integrator acts as a low-pass filter with a cutoff frequency of about 5.3 Hz. A compressive (logarithmic) nonlinearity is applied to the smoothed signal. The resolution of the system is limited by the addition of an internal noise. Finally, an optimal detector, with inputs from all monaural and binaural channels, is used as the decision device.

The model does not track perceived motion or predict spatial perception of the sound source, but rather looks at the energy in each EI channel (i.e., for a fixed ILD and ITD combination) in order to detect a signal. Diotic signals have no energy in the ILD=0, ITD=0 channel (perfect cancellation, internal noise is added later) and any interaural decorrelation will result in an increase in energy in this channel. Therefore, the addition of an antiphasic tone to a diotic noise (N_0S_π) will result in a much larger increase in energy than the addition of a homophasic tone (N_0S_0), demonstrating the classic BMLD (see e.g., Licklider, 1948; Hirsh, 1948).

The experimental conditions described in Sec. IV were simulated using the model. The simulation results are summarized as masking curves in Fig. 8, like those shown in Fig. 4(b). It is clear from a comparison of the human listeners' (open symbols) and the model's results (closed symbols) that the tuning described in Sec. IV is not captured by the model. The model does show a small peak at the signal frequency, but then the masking level increases with higher relative masker center frequencies, while the human listeners show a decrease in masking level with higher masker frequencies. The reason for the increase in masking in the model with high masker center frequencies is that the sliding integrator smooths out the interaural fluctuations from the masker, thereby removing any locations with good cancellation and increasing the energy at the output of the EI channels. Adding the diotic AM to the reference intervals further increases the energy so that there is less of a difference between the signal and reference intervals, making the discrimination task harder to perform. Part of this effect stems from the fact that the model was not designed to look at temporal differences and only compares the total energy at the output of the EI channels.

In order for this binaural model to be able to predict thresholds with fluctuating stimuli, it requires frequency selectivity in the processing of monaural and interaural level fluctuations. The monaural modulation filterbank (MFB) from Dau *et al.* (1997a) could be added to the peripheral channels, but then a question arises as to the sequence of model stages: Should the taps for the EI array come from before or after this filterbank? Two possible design concepts are shown in Fig. 9. The sequence of the stages would not be important except for the nonlinearities in both the monaural MFB and the EI process. The monaural MFB has a nonlinear reduction of modulation phase information for frequencies above 10 Hz. Without the modulation phase information, there would be no interaural differences with an AM_π signal, and the model would not be able to discriminate between AM_π and AM_0. If the EI inputs were to come from after the monaural modulation filters, but before this nonlinearity (left panel of Fig. 9), then the sharpness of tuning would be preserved through the output of the EI elements. That sharpness might be reduced to fit the measured data by adding additional noise and/or interaural differences in the processing, but the effect of this additional noise on other experiments would have to be investigated. Another option would be to take the EI inputs from before the monaural modulation filterbank (right panel of Fig. 9). In this manner, the interaural modulation phase differences would be preserved going into the binaural processor. However, a new model stage would then be required, namely a ILD modulation filterbank at the output of the EI system in addition to the two monaural amplitude modulation filterbanks. This filterbank would replace the sliding integrator from the original Breebaart model. The optimal sequence for the linear and nonlinear model stages should be investigated in further simulations.

VI. SUMMARY AND CONCLUSIONS

The first experiment showed that interaurally correlated and uncorrelated narrow-band noise carriers have a signifi-

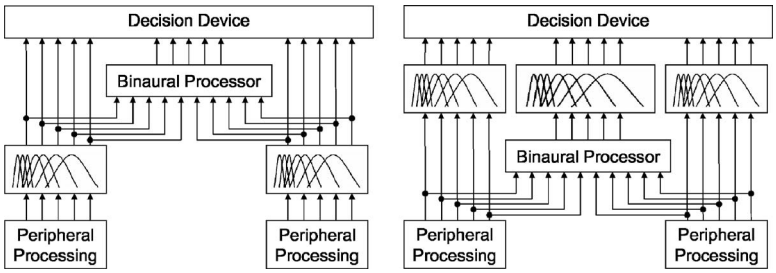

FIG. 9. Possible concepts for the inclusion of modulation frequency selectivity in binaural processing of ILD fluctuations. In the left panel, the modulation filterbank (MFB) from Dau *et al.* (1997a) is applied to each monaural peripheral channel before the input to the binaural processor. The binaural system would thereby inherit its ILD modulation frequency tuning from the monaural AM processing. The second concept, in the right panel, takes the inputs to the binaural processor from before the monaural MFBs. This then requires an additional binaural ILD modulation filterbank to add frequency selectivity in the processing of interaural level fluctuations.

cant effect on the discriminability of modulated ILDs (AM_π) from diotic AM (AM_0), particularly for modulation frequencies below the bandwidth of the carrier. This suggested that the binaural system shows broad bandpass modulation frequency tuning in processing of ILD fluctuations. A comparison of the results obtained with diotically modulated and unmodulated references underscored the importance of eliminating monaural cues in the design of binaural detection tasks because the signal detection will be based on monaural detection if the monaural cues are more salient than the binaural cues.

This modulation frequency tuning was further explored in the second experiment with AM_π discrimination in the presence of masking narrow-band noise modulators. The masking patterns also showed bandpass tuning, but with a broader tuning than that shown in similar monaural experiments (e.g., Ewert *et al.*, 2002).

An analysis with an existing binaural model (from Breebaart *et al.*, 2001a) showed that the model, which uses a low-pass filter to limit its temporal resolution in the processing of fluctuating interaural differences instead of a modulation filterbank, cannot predict the thresholds or the masking patterns measured with human listeners.

Further experiments should be performed to investigate the effect of diotic level fluctuations on the perception of ILD fluctuations through additional psychoacoustic tests as well as modeling. In addition, a binaural model should be developed that can predict the frequency selectivity shown here in the processing of interaural level fluctuations.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Richard Freyman, the associate editor, Dr. Wes Grantham, and three anonymous reviewers for their very helpful comments and suggestions on earlier revisions of the manuscript. This research was supported by a Ph.D. scholarship from the Technical University of Denmark.

[1]Low-noise is a bandpass noise for which the phase angles of the individual frequency components have been optimally selected to minimize the fourth moment of the signal, thereby minimizing the envelope fluctuations. Kohlrausch *et al.* (1997) produced low-noise noise using an iterative process of bandpass filtering the noise and normalizing the noise with its envelope. Each filtering step produces envelope fluctuations, and each normalization step produces a flat envelope, but a broader frequency spectrum. After many iterations, a noise is produced which has both a flat envelope and the desired bandwidth. The control experiment was performed with one listener, a 30-Hz-wide carrier, and a 4 Hz AM_π signal.

Bacon, S. P., and Grantham, D. W. (**1989**). "Modulation masking: Effects of modulation frequency, depth, and phase," J. Acoust. Soc. Am. **85**, 2575–2580.
Bernstein, L. R., and Trahiotis, C. (**1994**). "Detection of interaural delay in high-frequency sinusoidally amplitude-modulated tones, two-tone complexes, and bands of noise," J. Acoust. Soc. Am. **95**, 3561–3567.
Blauert, J. (**1972**). "On the lag of lateralization caused by interaural time and intensity differences," Audiology **11**, 265–270.
Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001a**). "Binaural processing model based on contralateral inhibition. I. Model structure," J. Acoust. Soc. Am. **110**, 1074–1088.
Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001b**). "Binaural processing model based on contralateral inhibition. II. Dependence on spectral parameters," J. Acoust. Soc. Am. **110**, 1089–1104.
Breebaart, J., van de Par, S., and Kohlrausch, A. (**2001c**). "Binaural processing model based on contralateral inhibition. III. Dependence on temporal parameters," J. Acoust. Soc. Am. **110**, 1105–1117.
Buell, T. N., Trahiotis, C., and Bernstein, L. R. (**1991**). "Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays," J. Acoust. Soc. Am. **90**, 3077–3085.
Dau, T. (**1996**). "Modeling auditory processing of amplitude modulation," Ph.D. thesis, Universität Oldenburg, Oldenburg, Germany.
Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," J. Acoust. Soc. Am. **102**, 2892–2905.
Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). "Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration," J. Acoust. Soc. Am. **102**, 2906–2919.
Dau, T., Püschel, D., and Kohlrausch, A. (**1996**). "A quantitative model of the 'effective' signal processing in the auditory system. I. Model structure," J. Acoust. Soc. Am. **99**, 3615–3622.
Durlach, N. I. (**1963**). "Equalization and cancellation theory of binaural masking-level differences," J. Acoust. Soc. Am. **35**, 1206–1218.
Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.
Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am. **112**, 2921–2931.
Feddersen, W. E., Sandel, T. T., Teas, D. C., and Jeffress, L. A. (**1957**). "Localization of high-frequency tones," J. Acoust. Soc. Am. **29**, 988–991.
Fleischer, H. (**1982**). "Modulationsschwellen von Schmalbandrauschen [Modulation thresholds of narrow-band noise]," Acustica **51**, 154–161.
Grantham, D. W. (**1982**). "Detectability of time-varying interaural correlation in narrow-band noise stimuli," J. Acoust. Soc. Am. **72**, 1178–1184.
Grantham, D. W. (**1984**). "Discrimination of dynamic interaural intensity differences," J. Acoust. Soc. Am. **76**, 71–76.
Grantham, D. W., and Bacon, S. P. (**1991**). "Binaural modulation masking," J. Acoust. Soc. Am. **89**, 1340–1349.
Grantham, D. W., and Wightman, F. L. (**1978**). "Detectability of varying interaural temporal differences," J. Acoust. Soc. Am. **63**, 511–523.
Griesinger, D. (**1997**). "The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces," Acust. Acta Acust.

83, 721–731.
Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex wave-forms," J. Acoust. Soc. Am. **55**, 84–90.
Hirsh, I. J. (**1948**). "The influence of interaural phase on interaural summation and inhibition," J. Acoust. Soc. Am. **20**, 536–544.
Houtgast, T. (**1989**). "Frequency selectivity in amplitude-modulation detection," J. Acoust. Soc. Am. **85**, 1676–1680.
Klumpp, R. G., and Eady, H. R. (**1956**). "Some measurements of interaural time difference thresholds," J. Acoust. Soc. Am. **28**, 859–860.
Kohlrausch, A., Fassel, R., and Dau, T. (**2000**). "The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers," J. Acoust. Soc. Am. **108**, 723–734.
Kohlrausch, A., Fassel, R., van der Heijden, M., Kortekaas, R., van de Par, S., Oxenham, A. J., and Püschel, D. (**1997**). "Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations," Acust. Acta Acust. **83**, 659–669.
Lawson, J. L., and Uhlenbeck, G. E. (**1950**). *Threshold Signals*, Radiation Laboratory Series Vol. **24** (McGraw-Hill, New York).
Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.
Licklider, J. C. R. (**1948**). "The influence of interaural phase relations upon the masking of speech by white noise," J. Acoust. Soc. Am. **20**, 150–159.
Mills, A. W. (**1960**). "Lateralization of high-frequency tones," J. Acoust. Soc. Am. **32**, 132–134.
Nuetzel, J. M., and Hafter, E. R. (**1981**). "Discrimination of interaural delays in complex waveforms: Spectral effects," J. Acoust. Soc. Am. **69**, 1112–1118.
Price, R. (**1955**). "A note on the envelope and phase-modulated components of narrow-band gaussian noise," IRE Trans. Inf. Theory **1**, 9–13.
Pumplin, J. (**1985**). "Low-noise noise," J. Acoust. Soc. Am. **78**, 100–104.
Schroeder, M. R. (**1981**). "Modulation transfer-functions: Definition and measurement," Acustica **49**, 179–182.
Sheft, S., and Yost, W. A. (**2007**). "Discrimination of starting phase with sinusoidal envelope modulation," J. Acoust. Soc. Am. **121**, EL84–EL89.
Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (**2005**). "Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers," J. Acoust. Soc. Am. **118**, 2507–2518.
Terhardt, E. (**1968**). "Über die durch amplitudenmodulierte Sinustöne hervorgerufene Hörempfindung. [The auditory sensation produced by amplitude modulated tones]," Acustica **20**, 210–214.
Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.
Witton, C., Green, G. G., Rees, A., and Henning, G. B. (**2000**). "Monaural and binaural detection of sinusoidal phase modulation of a 500-Hz tone," J. Acoust. Soc. Am. **108**, 1826–1833.