

Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers

Mark A. Stellmack,^{a)} Neal F. Viemeister, and Andrew J. Byrne
Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 28 September 2004; revised 22 June 2005; accepted 18 July 2005)

Temporal modulation transfer functions (TMTFs) were measured for detection of monaural sinusoidal amplitude modulation and dynamically varying interaural level differences for a single set of listeners. For the interaural TMTFs, thresholds are the modulation depths at which listeners can just discriminate interaural envelope-phase differences of 0 and 180°. A 5-kHz pure tone and narrowband noises, 30- and 300-Hz wide centered at 5 kHz, were used as carriers. In the interaural conditions, the noise carriers were either diotic or interaurally uncorrelated. The interaural TMTFs with tonal and diotic noise carriers exhibited a low-pass characteristic but the cutoff frequencies changed nonmonotonically with increasing bandwidth. The interaural TMTFs for the tonal carrier began rolling off approximately a half-octave lower than the tonal monaural TMTF (~80 Hz vs ~120 Hz). Monaural TMTFs obtained with noise carriers showed effects attributable to masking of the signal modulation by intrinsic fluctuations of the carrier. In the interaural task with dichotic noise carriers, similar masking due to the interaural carrier fluctuations was observed. Although the mechanisms responsible for differences between the monaural and interaural TMTFs are unknown, the lower binaural TMTF cutoff frequency suggests that binaural processing exhibits greater temporal limitation than monaural processing. © 2005 Acoustical Society of America. [DOI: 10.1121/1.2032057]

PACS number(s): 43.66.Pn, 43.66.Mk [AK]

Pages: 2507–2518

I. INTRODUCTION

Comparisons of intensity across the ears and comparisons of intensity over time within one ear are performed physiologically by different mechanisms. Neural circuits in the lateral superior olive are configured such that they effectively compare intensity information arriving simultaneously at the two ears (Yin, 2002). The result is neurons that are “tuned” to interaural level differences (ILDs). Perceptually, for stimuli presented to the two ears over headphones that differ only in interaural level, an intracranial image is typically perceived at some lateral position determined by the ILD. Thus, binaural comparisons of intensity result in a spatial component to the internal representation of the stimulus. In contrast, monaural comparisons of intensity over time would seem to depend upon the intensity of each stimulus being encoded in some way (perhaps in terms of loudness) and retention of those representations in memory for subsequent comparison. Stellmack *et al.* (2004) presented data indicating that, despite these physiological and perceptual differences, the auditory system is nearly equally sensitive to differences in intensity for a 4-kHz pure tone and broadband noise when those differences are presented over time to one ear or simultaneously to both ears. In addition, both monaural and binaural comparisons of intensity were shown to exhibit Weber’s law for broadband noise and the near miss for a 4-kHz pure tone, likely reflecting the effects of common stages of processing preceding monaural and binaural comparison.

The above-described conditions involved monaural and binaural comparisons of intensities that remained constant for the duration of the individual stimuli. The present paper examines monaural and binaural processing of dynamically varying stimulus intensity. It has been shown that the auditory system exhibits a decreasing ability to resolve the fluctuations associated with sinusoidal amplitude modulation (SAM) with increasing modulation frequency and that this limitation is not due to the reduction of effective modulation depth as a result of critical-band filtering of the stimulus (e.g., Viemeister, 1979; Kohlrausch *et al.*, 2000). It also has been shown that binaural processing exhibits a limitation in temporal resolution that is not attributable to critical-band filtering (Bernstein and Trahiotis, 1994; 2005).

Temporal limitations in monaural and binaural processing are often modeled as integration of the relevant parameter of the stimulus (e.g., intensity in the monaural case and interaural parameters in the binaural case) within a temporal window that is limited by some minimum integration time. Because of the minimum integration time, rapid fluctuations in the parameter of interest are effectively “smoothed over” or low-pass filtered by the system, producing a limitation on the temporal resolution of the system. This limitation in temporal resolution is often characterized in terms of the cutoff frequency or time constant of a low-pass filter or in terms of the minimum integration time of a temporal window that produces an equivalent limitation on temporal resolution.

A number of studies have estimated cutoff frequencies and time constants or minimum integration times for monaural and binaural processing of dynamically varying stimulus parameters. Generally, these studies have shown that binaural processing exhibits lower temporal resolution than mon-

^{a)}Electronic-mail: stell006@umn.edu

aural processing (e.g., Grantham and Wightman, 1979; Kollmeier and Gilkey, 1990; Culling and Summerfield, 1998). For example, Kollmeier and Gilkey (1990) measured binaural temporal resolution through binaural forward- and backward-masking tasks in which listeners detected an interaurally out-of-phase signal (S_π) in the presence of a concurrent interaurally in-phase noise masker (N_0), where the in-phase masker was preceded or followed by interaurally phase-inverted noise (N_π). Thresholds were measured as a function of the time between the signal presentation and the change in the masker configuration. Time constants of a two-sided exponential window that produced best fits to the data ranged from 16.1 to 42.7 ms (across all listeners and conditions). In contrast, estimated time constants of a temporal window that predicted data in comparable monaural forward- and backward-masking tasks ranged from 5.9 to 14.7 ms, indicating better monaural temporal resolution. In an extension of the Kollmeier and Gilkey (1990) task, Culling and Summerfield (1998) measured thresholds for an S_π signal in the presence of N_0 noise that was both preceded and followed by N_π noise. In considering a number of different temporal windows, Culling and Summerfield (1998) found that an asymmetrical Gaussian window fit their data best. Best-fitting time constants for the forward lobe ranged from 11.7 to 28.7 ms for the forward lobe and from 33.3 to 61.1 ms for the rearward lobe, reflecting more masking of the S_π signal by the leading N_π fringe than the trailing N_π fringe. Grantham and Wightman (1979), in a task in which listeners detected an S_π signal in the presence of a masker whose interaural configuration varied sinusoidally between N_0 and N_π , estimated binaural time constants that ranged from 44 to 243 ms across listeners and conditions. In the present experiments, the ability to detect periodically time-varying ILDs is compared to monaural detection of SAM in order to compare limitations on temporal resolution in the two situations. A single set of listeners was run in analogous monaural and binaural tasks that allow for a direct comparison between monaural and binaural performance.

A. Monaural TMTF

In a typical monaural modulation detection task, the listener attempts to detect sinusoidal fluctuations in amplitude that are present in the signal interval but not in the nonsignal interval. In this case, the limits of monaural temporal resolution can be described by the temporal modulation transfer function (TMTF), which plots threshold modulation depth as a function of modulation frequency (Viemeister, 1979). For broadband noise carriers, thresholds are relatively constant across a range of low modulation frequencies, while above some cutoff modulation frequency thresholds increase with increasing modulation frequency, in effect describing a low-pass characteristic for modulation processing. The form of the TMTF generally is interpreted as indicating that the fluctuations of the stimulus envelope are encoded by the auditory system with less precision as modulation frequency increases. That is, the representation of the envelope within the auditory system becomes “smoothed over” at higher modulation frequencies, such that an increase in modulation depth

is required in order for the modulation to be detected. The cutoff modulation frequency and rate of increase of thresholds (representing a decrease in sensitivity) with increasing modulation frequency sometimes are used to describe the auditory system’s low-pass characteristic for modulation. The form of the TMTF measured with puretone carriers differs from that for broadband noise carriers in that the puretone TMTF typically has a higher cutoff modulation frequency (Dau *et al.*, 1999; Kohlrausch *et al.*, 2000). For example, in the TMTFs measured with a high-frequency pure-tone carrier by Kohlrausch *et al.* (2000), thresholds are relatively constant for modulation frequencies below about 130 Hz. As modulation frequency increases above 130 Hz, thresholds first increase as temporal resolution diminishes and then decrease as the modulation sidebands become resolved at the auditory periphery (see the following). Because the TMTFs measured with noise and tonal carriers exhibit different cutoff frequencies, one must consider the potential advantages and disadvantages of using each type of carrier to measure the monaural TMTF.

When a pure tone is sinusoidally amplitude modulated, two spectral sidebands are introduced, one above and one below the carrier frequency, with frequency spacings equal to the modulation frequency. When the modulation frequency is low, the sidebands are unresolved at the auditory periphery and the listener performs the modulation-detection task by detecting the fluctuating stimulus envelope. When the modulation frequency becomes sufficiently high, the sidebands become increasingly resolved and the listener simply detects the sidebands rather than the modulation *per se* (e.g., Kohlrausch *et al.*, 2000). This leads to sharply decreasing thresholds at high modulation frequencies and can make it difficult to obtain an accurate assessment of the temporal resolution of the auditory system.

In contrast, the long-term spectrum of a broadband noise carrier remains unchanged when it is sinusoidally amplitude modulated, so this provides a potential means for studying auditory temporal resolution at high modulation frequencies. However, the envelope of a broadband noise itself fluctuates in amplitude. These intrinsic envelope fluctuations, in turn, produce envelope fluctuations at the outputs of peripheral auditory filters as a result of their bandpass filtering of the broadband noise. It has been suggested that these intrinsic fluctuations in the noise carrier mask the sinusoidal modulation to be detected (Dau *et al.*, 1999). In order for this type of explanation to account for the data, one must assume some type of frequency selectivity in the modulation domain, for example, processing of the peripherally filtered signal by modulation frequency-selective channels, a so-called modulation filterbank (Dau *et al.*, 1997). In this scheme, when both the signal modulation and intrinsic fluctuations of the carrier are passed by a common modulation filter, the signal modulation is less detectable resulting in higher thresholds. Consistent with that notion, the monaural TMTF for a broadband noise carrier starts to increase at lower modulation rates and with a different slope than the TMTF for a pure-tone carrier (Kohlrausch *et al.*, 2000). Also consistent with the concept of processing by a modulation filterbank, TMTFs for narrowband noise exhibit a bandpass or high-pass character-

istic depending upon the bandwidth of the carrier, which approximately determines the frequency range of the intrinsic carrier fluctuations (Dau *et al.*, 1999). If the fact that thresholds in the TMTF for broadband noise increase at a lower frequency is in part a result of masking of the signal modulation by the intrinsic fluctuations of the carrier, this produces a deceptively low estimate of the temporal resolution of the auditory system. Therefore, it seems that the best stimulus for use in assessing monaural temporal resolution is a pure-tone carrier (to avoid intrinsic carrier fluctuations) that is very high in frequency (to minimize resolution of the sidebands with increasing modulation frequency).

B. Interaural TMTF

What is the appropriate stimulus for assessing the temporal resolution of the binaural system? Consider a dichotic stimulus in which a sinusoidally amplitude-modulated pure tone is presented to each ear, with the envelopes at the two ears 180° out-of-phase with one another. The resulting stimulus is, in effect, a pure tone for which the ILD oscillates about zero at a rate equal to the monaural modulation frequency in each ear, as shown in the right column at the top of Fig. 1, in which the instantaneous ILD functions were defined as

$$\text{ILD}(t) = 20 \log_{10} \left(\frac{E_L(t)}{E_R(t)} \right), \quad (1)$$

where $E_L(t)$ and $E_R(t)$ are the envelopes of the left- and right-ear signals, respectively. [The function $\text{ILD}(t)$ is nearly sinusoidal for the threshold modulation depths measured in this paper, while the form of the function approaches that of a triangle wave for larger monaural modulation depths. See Fig. 2.] At low modulation frequencies, this stimulus typically produces the percept of a tone that moves slowly from side to side within the listener's head. The extent of the perceived lateral movement is determined by the peak interaural level difference that is attained by the stimulus, which increases with increasing monaural modulation depth. At higher modulation frequencies, the stimulus is no longer perceived as a discrete, moving image but instead as a diffuse or broadened image (when the envelopes at the two ears are 180° out-of-phase with one another), with the degree of diffuseness directly related to the monaural modulation depth. Variation of the stimulus ILD over time will be called "interaural modulation" in this paper.

In his study of interaural modulation detection, Grantham (1984) sought to eliminate "diffuseness" as a cue by using an interaurally uncorrelated noise as the carrier. When the carrier is interaurally uncorrelated, the instantaneous interaural level difference of the noise carrier fluctuates randomly about zero (Fig. 1, lower panels) which itself produces a perception of diffuseness. [In Fig. 1, the instantaneous ILD functions in the right-hand panels were computed from the Hilbert envelopes of the corresponding noise wave forms using Eq. (1).] In this case, the listener must detect sinusoidal interaural modulation in the presence of random interaural fluctuations intrinsic to the interaurally uncorrelated noise carrier. Note that this is analogous to the

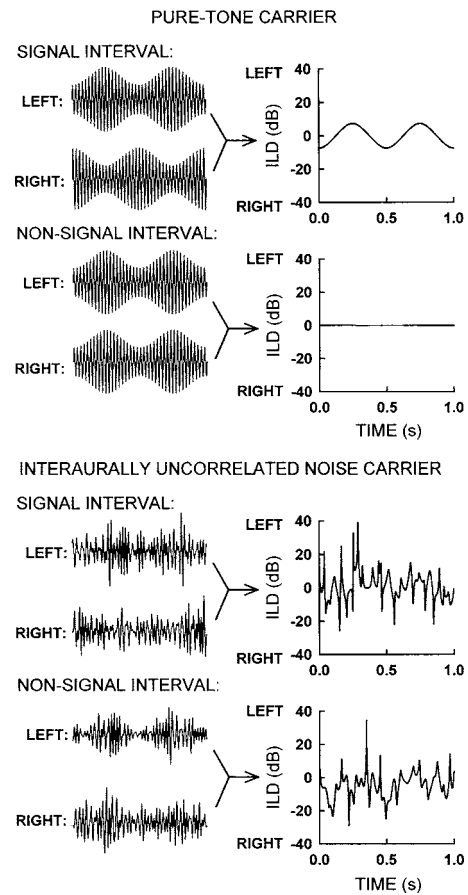


FIG. 1. The pressure wave forms presented to the left and right ears (left column) and instantaneous interaural level difference (ILD, right column) of stimulus envelopes in signal and nonsignal intervals of an interaural modulation detection task. In the upper set of panels, the carrier is a pure tone while in the lower panels the carrier is a 30-Hz-wide interaurally uncorrelated band of noise. The depth of the sinusoidal amplitude modulation of each stimulus is the same in the upper and lower panels ($m=0.4$). The instantaneous ILD functions in the lower panels were computed from the Hilbert envelopes of the noise wave forms using Eq. (1).

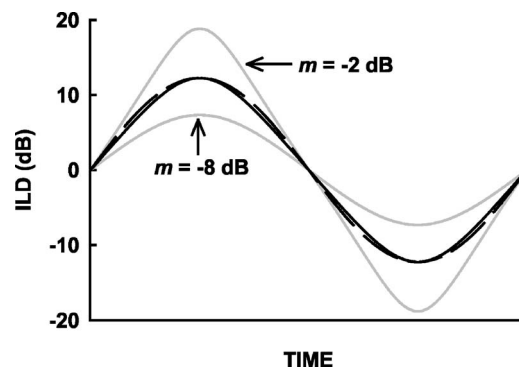


FIG. 2. One cycle of $\text{ILD}(t)$ (as in Fig. 1) computed using Eq. (1) for several SAM pure tones with various monaural modulation depths. The solid black line is $\text{ILD}(t)$ for a stimulus with $m = -4.33$ dB, the largest threshold measured in this experiment. The dashed black line is a sinusoid with the same peak amplitude as the solid black line. The largest instantaneous difference between the solid and dashed black lines is 0.71 dB. The gray lines represent $\text{ILD}(t)$ for the values of m indicated in the figure. It can be seen that $\text{ILD}(t)$ is nearly sinusoidal for the range of threshold m measured in this experiment.

situation in the monaural modulation detection task with a broadband noise carrier in which the listener must detect sinusoidal amplitude modulation in the presence of random fluctuations of amplitude. It is possible that the intrinsic interaural fluctuations of the interaurally uncorrelated carrier might interfere with detection of the sinusoidal interaural modulation. Therefore, it seems that the best stimulus for measuring an interaural TMTF that is analogous to the monaural TMTF is a diotic carrier, for which there are no intrinsic interaural fluctuations that might interfere with detection of the signal interaural modulation.

In the present experiment, monaural and interaural modulation detection is compared for narrowband noise carriers (30- and 300-Hz wide centered at 5 kHz) for which intrinsic fluctuations of the carrier might be expected to interfere with detection of the target sinusoidal modulation. In separate conditions, interaural modulation detection thresholds are measured for diotic and dichotic (interaurally uncorrelated) noise carriers. In the diotic case, there are no intrinsic interaural fluctuations in the carrier that might interfere with the interaural modulation to be detected. The dichotic conditions are included to assess the effects on the interaural TMTF of degrading the “image-width” cue. Monaural and interaural modulation detection also is measured for a 5-kHz pure-tone carrier for which there likewise are no intrinsic fluctuations in the monaural or interaural modulation functions. Furthermore, given the 5-kHz carrier, the sidebands of the modulated stimulus are high enough in frequency so that interaural delays of the resolved sidebands cannot serve as a cue for detection in the binaural condition (Klumpp and Eady, 1956). Such a cue can complicate the interpretation of interaural TMTFs measured with low-frequency carriers. For the narrowband noise carriers, when the sinusoidal modulators are interaurally 180° out of phase, the phases of the fine structures of the modulation sidebands are inverted relative to one another but the envelopes of the modulation sidebands are identical at the two ears. Thus, even if the modulation sidebands of the narrowband noise carrier are resolved, they carry no interaural envelope cues to the detection of the interaural modulation in the signal interval.

The primary focus of the present paper is a comparison of monaural and binaural temporal resolution as reflected by the monaural and interaural TMTFs. An additional issue that will be addressed is the influence of intrinsic fluctuations in monaural intensity or interaural intensity differences on the forms of the monaural and interaural TMTFs.

II. METHODS

A. Stimuli and procedure

In all (monaural and interaural) conditions, a three-interval, three-alternative forced-choice task consisting of one signal and two nonsignal intervals was used. In the monaural conditions, the nonsignal intervals consisted of unmodulated carriers and the signal interval contained a sinusoidally amplitude-modulated carrier, with signals and nonsignals presented only to the left ear. The starting phase of the envelope in the signal interval was chosen randomly from trial to trial.

In the interaural modulation-detection task, stimuli were presented to both ears and were sinusoidally amplitude-modulated in all three intervals of every trial. In the signal interval, the interaural phase difference of the sinusoidal modulator was 180° while the interaural phase difference of the sinusoidal modulator was 0° in the nonsignal intervals. The absolute starting phase of the modulator was randomized between intervals (while maintaining the appropriate interaural phase difference) to eliminate possible monaural cues associated with changing the phase of one channel in the signal interval.

The monaural and interaural tasks were performed with a number of different carriers. For the monaural tasks, the carriers were a 5-kHz pure tone and 30- and 300-Hz-wide bands of noise centered at 5 kHz. These bandwidths were chosen to permit comparison to the data of Dau *et al.* (1997). In the interaural tasks, the carriers were a 5-kHz pure tone, 30- and 300-Hz-wide diotic bands of noise, and 30- and 300-Hz-wide interaurally uncorrelated bands of noise, with all noise bands centered at 5 kHz. The pure-tone carrier was interaurally in-phase in the interaural condition. Independent samples of narrowband noise were generated for the left- and right-ear stimuli of the interaurally uncorrelated carriers. For all of the noise carrier conditions, different samples of noise were generated for each interval of each trial (that is, the carriers were not “frozen” across intervals).

The monaural and interaural conditions were run in the presence of continuous, low-pass noise with a cutoff frequency of 1.3 kHz designed to mask potential distortion products at the modulation frequency (e.g., Nuetzel and Hafter, 1976; Wiegand and Patterson, 1999). In addition, the monaural TMTF was measured without the low-pass masker in order to facilitate comparison to the data of Kohlrausch *et al.* (2000), who also measured thresholds without a low-pass masker.

In all conditions, modulation depth was varied adaptively using a two-down, one-up adaptive tracking rule that estimated the 70.7% correct point on the psychometric function (Levitt, 1971). Modulation depth was varied in the signal and nonsignal intervals in the interaural modulation detection task. (Recall that varying the monaural modulation depth in the interaural modulation condition has the effect of varying the peak interaural level difference attained by the dichotic stimulus.) The initial step size was set to 2 dB (in units of $20 \log m$, where m is the modulation index) and was reduced to 1 dB after four reversals. A block of trials was terminated after a total of 12 reversals, and the modulation depths at which the final eight reversals occurred (in dB) were averaged to produce a threshold estimate. If the adaptive procedure attempted to track above a modulation depth of 0 dB ($m=1$), the block of trials was terminated immediately. If the adaptive track was terminated prematurely in two consecutive runs, the threshold was considered to be unmeasurable. When possible to do so, the initial modulation depth was set approximately 8 dB above the expected threshold modulation depth based on prior experience and the performance of the listeners. The highest initial modulation depth used was -5 dB to allow listeners to make at least two incorrect responses without prematurely terminating the adap-

tive run because of overmodulation. Four blocks of trials were run in each condition and the four threshold estimates (in dB) were averaged to obtain the final threshold estimate for that condition.

The 5-kHz pure-tone carrier was presented at a level of 75 dB SPL at each ear. The 30- and 300-Hz-wide bands of noise were presented at spectrum levels of 50 and 40 dB, respectively, measured at 1-kHz for unfiltered stimuli. The temporal characteristics of the stimuli were based upon those used by Kohlrausch *et al.* (2000) although there were minor differences in the specific parameters. Each interval was 1 s in duration (the intervals of Kohlrausch *et al.* were 800 ms in duration), including 150-ms raised-cosine on-off ramps. The intervals were temporally contiguous with no silence between them, the on-off ramps serving to demarcate the intervals. Thresholds were measured for modulation frequencies (f_m) ranging from 4 to 650 Hz with all but the lowest f_m identical to those used by Kohlrausch *et al.* (2000).

Listeners were given a list of computer files associated with each experimental condition and were allowed to initiate adaptive runs themselves. They were given instructions to complete two adaptive runs for each f_m in an order of their choosing and then two additional runs in a different order, again of their choosing, and the experimenters confirmed that these instructions were followed. The instructions to complete the adaptive runs in various orders was intended to control for potential order effects on thresholds measured across f_m 's. As a result, listeners in many cases performed adaptive runs for the f_m 's within an experimental condition in different orders, for example, listeners may have run the f_m 's from low to high for the first two runs then high to low for the second two, or a listener might have chosen to proceed through the f_m 's in a less systematic way. All f_m 's for a particular condition were run before a different condition was run. Listeners ran the main conditions of the experiment in the following order: monaural tonal without masker, monaural tonal with masker, interaural tonal, monaural noise (both bandwidths), interaural with correlated noise (both bandwidths), and interaural with uncorrelated noise (both bandwidths). With regard to the conditions with noise carriers, listeners ran the 30- and 300-Hz bandwidth conditions in different orders.

To compensate for the increase in intensity resulting from amplitude modulation and to limit the usefulness of changes in overall intensity as a cue (a potential problem in the monaural modulation detection tasks), all modulated wave forms in all conditions were scaled by a factor of $(1 + m^2/2)^{-0.5}$. Although intensity changes associated with modulation did not provide a cue in the binaural listening tasks because all intervals were modulated, the modulated wave forms in all three intervals of the binaural conditions also were rescaled by the above-noted factor for consistency.

Each block of trials was initiated by the listener. On each trial, a "ready" light flashed on the computer screen for 250 ms followed by a 300-ms pause after which a trial was presented. A separate marker on the screen was illuminated during each stimulus interval. Listeners entered their responses on the computer keyboard at which time the correct answer was indicated by once again illuminating the appropriate in-

terval marker on the screen. Listeners were run in 2-h sessions, during which approximately 12–14 blocks of trials were run, until all stimulus conditions were completed.

B. Apparatus

Signals and nonsignals were generated and presented via MATLAB (Math Works) on a PC equipped with a high-quality, 24-bit sound card (Echo Audio Gina) at a sampling rate of 44.1 kHz. Each noise carrier was generated digitally in the frequency domain with maximally steep roll-offs at the cut-off frequencies and was converted to the time domain via a 44 100-point inverse fast Fourier transform (FFT) that yielded a 1-s sample of noise. The low-pass masker used in the interaural conditions was dichotic Gaussian noise produced continuously at each ear by separate noise generators (General Radio 1381). The noise at each ear was low-pass filtered by a programmable filter (Kemo VBF/25) with a cut-off of 1.3 kHz (135-dB/octave rolloff). The noise had a spectrum level of 40 dB SPL measured at 1 kHz after filtering. When the continuous noise masker was presented in the monaural conditions, it was presented to only the left ear. Stimuli were mixed with the noise maskers and presented over Sony MDR-V6 stereo headphones to listeners seated in an IAC sound-attenuating chamber.

C. Subjects

The three listeners consisted of the first and third authors (S1 and S2, respectively) and one female undergraduate student from the University of Minnesota (S3) who was paid to participate in the study. All listeners had pure-tone thresholds of 15 dB HL or better at octave frequencies from 250 to 8000 Hz. S1 and S2 had extensive experience in similar listening tasks and received no training prior to performing the present tasks. S3 was allowed to practice in a variety of the monaural and binaural conditions over the course of 2 weeks (approximately 12 h of practice) after which time she appeared to have reached asymptotic performance.

III. RESULTS

The pattern of results was very similar in all but one condition for the three listeners, so only the mean data are shown in Fig. 3. (The exception, the interaural TMTFs measured with a diotic 300-Hz-wide noise carrier, will be discussed in Sec. IV A.) In the left column of Fig. 3, the data are shown in the top panel for the tonal carrier, the middle panel for the 30-Hz-wide noise, and in the bottom panel for the 300-Hz-wide noise. In the right column of Fig. 3, the data from the left column are replotted such that the monaural data are in the top panel and the interaural data are in the bottom panel, with the same symbols representing the same conditions across the two columns. In each panel, threshold modulation depth is plotted as a function of f_m , with the ordinate inverted such that larger modulation depths are toward the bottom of the panel. Dashed portions of the interaural functions indicate that thresholds could not be measured above a particular modulation frequency. Some general

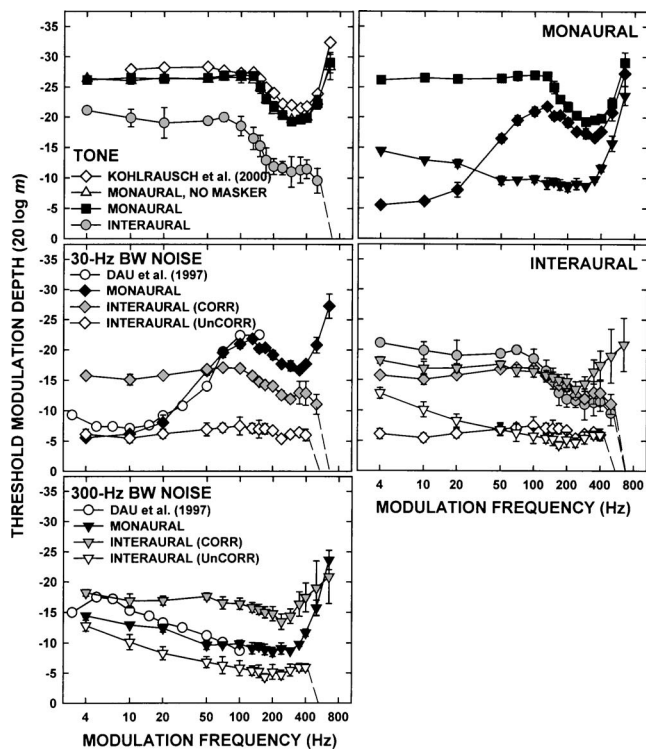


FIG. 3. Monaural and interaural TMTFs showing threshold modulation depth (in dB) as a function of modulation frequency averaged across three listeners. In the left column, TMTFs for different carriers are represented in each panel: top panel, 5-kHz pure-tone carrier; middle panel, 30-Hz-wide band of noise; bottom panel, 300-Hz-wide band of noise. The two bands of noise were centered at 5 kHz. The different symbol types represent monaural and interaural TMTFs, as indicated in the legend. Error bars represent standard errors of the mean computed across three listeners and are in many cases smaller than the symbols. In the top panel, the monaural, no-masker TMTF is almost completely obscured by the monaural TMTF (with masker). The open diamonds in the top panel represent the monaural TMTF for a 5-kHz pure-tone carrier reported by Kohlrausch *et al.* (2000). The open circles in the middle and bottom panels represent monaural TMTFs for 31- and 314-Hz-wide bands of noise, both centered at 5 kHz, reported by Dau *et al.* (1997). In the right column, in separate panels, the monaural and interaural TMTFs are replotted from the left column. The same symbol type represents the same data in each column.

features of the TMTFs will be described in this section, while specific parameter estimates are described in Sec. IV A.

Referring to the top left panel of Fig. 3, the monaural data gathered in the presence of a low-pass masker (filled squares) fall almost directly on top of the monaural data gathered without a masker (open triangles), indicating that the continuous low-pass masker had no effect on the monaural pure-tone TMTFs. (Note that all remaining monaural and binaural thresholds were measured in the presence of the low-pass masker.) Rickert and Viemeister (1998) measured TMTFs for a 2-kHz pure tone carrier with and without a broadband noise masker designed to mask distortion products and reduce off-frequency listening. (Figures showing those TMTFs were published by Viemeister *et al.*, 2002.) Rickert and Viemeister found that thresholds were elevated in the presence of noise, but the TMTF measured with noise was similar in form to that measured without noise. They concluded that the form of the TMTF is not dependent upon the detection of distortion products.

The form of the tonal monaural TMTF measured here is very similar to that reported by Kohlrausch *et al.* (2000) for a 5-kHz tone (open diamonds) despite some procedural differences (e.g., Kohlrausch *et al.* used 800-ms stimuli and varied the carrier frequency and level). Monaural thresholds are nearly constant up to about $f_m=130$ Hz, above which thresholds begin to increase. Thresholds decrease as f_m increases above 290 Hz, presumably reflecting detection of the increasingly resolved modulation sidebands.

In the middle and lower panels in the left column of Fig. 3, the monaural TMTFs for narrowband noise carriers (filled symbols) were similar in form to those reported by Dau *et al.* (1997) for f_m up to about 100–150 Hz (open circles). These similarities exist despite the fact that Dau *et al.* used slightly different bandwidths (31 and 314 Hz), shorter-duration stimuli (300 ms), and they bandpass filtered their wider-bandwidth stimuli back down to 314 Hz after modulation to limit spectral cues to the presence of modulation. Bandpass filtering after modulation reduces the effective modulation depth by an amount dependent upon modulation frequency (for example, for the 314-Hz-wide stimuli, by just over 3 dB for $f_m=100$ Hz and by about 1.5 dB for $f_m=50$ Hz).

The effects of intrinsic fluctuations in the narrowband noise carriers on the form of the TMTF can be examined by comparing the monaural TMTFs for the tone and noises (top right panel of Fig. 3). For the 30-Hz-wide carrier, thresholds are drastically elevated at low f_m 's relative to the pure-tone TMTF. As described in Sec. I, this reflects the fact that the power in the envelope spectrum of the unmodulated carrier is concentrated below the frequency corresponding to its bandwidth and that these intrinsic low-frequency fluctuations in the carrier interfere with detection of low rates of sinusoidal modulation (assuming a modulation frequency-selective mechanism with limited spectral resolution such as a modulation filterbank; Dau *et al.*, 1999). Relative to the tonal TMTF, thresholds for the 300-Hz-wide carrier are elevated across the entire range of f_m used here, reflecting the fact that energy in the envelope of the unmodulated 300-Hz-wide carrier extends over much of that range of envelope frequency.

Referring to the interaural TMTFs for the diotic carriers (the pure tone and two interaurally correlated carriers; see the shaded symbols in the lower right panel of Fig. 3), the roll-off rates seem to decrease with increasing carrier bandwidth but it appears that all show a low-pass characteristic with roll-offs beginning at modulation frequencies above about 50–100 Hz. The pure-tone and 30-Hz TMTFs roll off to a plateau from around $f_m=200$ –400 Hz, above which thresholds increase abruptly to the extent that thresholds were unmeasurable above $f_m=500$ Hz. Mean thresholds for the 300-Hz-wide, correlated noise carrier (shaded triangles) decrease substantially above $f_m=240$ Hz. This decrease in thresholds is due primarily to the data of one listener and will be discussed in Sec. IV A.

The TMTFs for the interaurally uncorrelated noise carriers (open symbols in the lower right panel of Fig. 3) are more irregular in form with a less obvious low-pass characteristic than those for the diotic carriers. Overall thresholds with uncorrelated carriers are higher than those for the diotic carriers. For the uncorrelated-carrier interaural TMTFs,

TABLE I. For each set of data listed in the first column, the 3-dB-down cutoff frequency (f_c) is shown for the best-fitting, low-pass Butterworth filter of first order (second column) and second order (third column). Fits were performed on all data of each TMTF up to the modulation frequency at which the slope changed from positive to negative. Parameters of the best-fitting functions consisting of two line segments are shown in the remaining columns. Fit 1 was performed on all data of each TMTF up to the modulation frequency at which the slope changed from positive to negative. Fit 2 was performed on one less data point than Fit 1 (the data point for the highest modulation frequency was excluded). The breakpoint for each fit is the frequency at the point of intersection of the two line segments, the 3-dB cutoff is the frequency at which the function is 3 dB down from its maximum, and the roll-off is the slope of the line above the breakpoint. The percentage in parentheses beneath each f_c and breakpoint is the percentage of variance accounted for by each fitted function [computed as shown in Eq. (2)].

	First-order Butterworth f_c (3 dB down)	Second-order Butterworth f_c (3 dB down)	2 line segments—Fit 1			2 line segments—Fit 2		
			Breakpoint (Hz)	3-dB cutoff (Hz)	roll-off (dB/oct)	Breakpoint (Hz)	3-dB cutoff (Hz)	roll-off (dB/oct)
Kohlrausch <i>et al.</i> (2000)	160 Hz (95%)	229 Hz (81%)	104 (97%)	177	3.91	114 (97%)	178	4.65
Monaural tone	144 Hz (84%)	184 Hz (92%)	116 (99%)	166	5.79	121 (98%)	167	6.60
Interaural tone	93 Hz (89%)	126 Hz (96%)	79 (96%)	117	5.21	85 (96%)	120	6.02
Interaural noise (30 Hz)	225 Hz (78%)	241 Hz (80%)	118 (90%)	224	3.24	113 (86%)	222	2.59
Interaural noise (300 Hz)	203 Hz (91%)	217 Hz (80%)	59 (89%)	210	1.63	48 (87%)	258	1.25

thresholds for $f_m=50$ -400 Hz are within a few dB of one another. Below $f_m=50$ Hz, thresholds for the 30-Hz-wide uncorrelated noise are higher than those for the 300-Hz-wide uncorrelated noise, as is the case for the monaural TMTFs. This indicates that intrinsic interaural fluctuations in the uncorrelated noise may affect the interaural TMTF in a way that is similar to the effect of intrinsic monaural fluctuations on the monaural TMTF. This will be discussed further in Sec. IV B.

IV. DISCUSSION

A. Monaural versus interaural TMTF

This section will focus on a comparison of monaural and interaural TMTFs measured with carriers that introduce no intrinsic fluctuations in the monaural or binaural domain of the signal modulation, that is, the monaural tonal TMTF and the interaural TMTFs measured with tonal or diotic noise carriers. In Fig. 3, it appears that the monaural tonal TMTF and the interaural TMTFs measured with diotic carriers show a low-pass characteristic. Thresholds are nearly constant over a broad range of low f_m 's and increase (reflecting attenuation of the modulation by the auditory system) above some cutoff frequency. As f_m increases further, thresholds eventually begin to decrease, changing the slope of the TMTF from positive to negative. In order to quantify the cutoff frequencies of the TMTFs, a process of fitting various functions to the data was undertaken. In the fits described in the following, unless noted otherwise, functions were fit only to the portion of the TMTFs up to the modulation frequency where the change from positive to negative slope occurred (i.e., only to the low-pass portion of the TMTFs).¹

As a first attempt to quantify their cutoff frequencies, each TMTF was fit with the inverted frequency response of a low-pass Butterworth filter. (Recall that the TMTFs are plot-

ted here on inverted ordinates.) The MATLAB Optimization Toolbox function “lsqnonlin” was used to estimate the cutoff frequency, order, and gain of the filter such that the mean-squared error between the inverted frequency response of the filter and the data was minimized. For each fit, predicted threshold values were computed and the percentage of variance in the actual data that was accounted for by the best-fitting filter was calculated as follows:

$$100 \times \left(1 - \frac{\sum (O_i - P_i)^2}{\sum (O_i - \bar{O})^2} \right), \quad (2)$$

where O_i and P_i are the observed and predicted threshold values, respectively, and \bar{O} is the mean of the observed threshold values. It was found that the best-fitting filters had orders of either 1 or 2, that is, neither order yielded best-fitting filters to all the data. This was due to the different apparent roll-offs of the TMTFs. In Table I, the cutoff frequency and percent of variance accounted for by filters of order 1 and 2 for each TMTF are shown.

In order to describe more precisely the differences between roll-off rates of the TMTFs, it was decided that the TMTFs would be fit with two straight lines, one with slope zero and a second with a positive slope. It was found that the greatest percentage of variance in the observed data could be accounted for in most cases by the straight-line fits, therefore the remainder of the discussion will refer only to the results of those fits. The fits were determined via least-squares linear regression for various groupings of the data within each TMTF to find the two lines that produced the maximum percentage of variance accounted for as computed in Eq. (2). The breakpoint (point of intersection of the two best-fitting straight lines), roll-off rate (positive slope), and percentage of variance accounted for by each fit are shown in the fourth

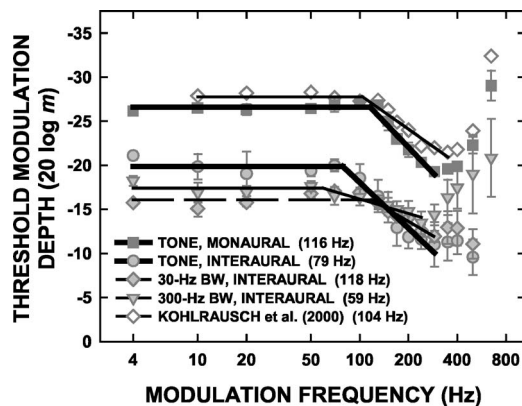


FIG. 4. The monaural tonal TMTF and the interaural TMTFs measured with a pure tone and diotic noise carriers fitted with functions consisting of two straight lines. The data are replotted from Fig. 3. The data points have been grayed out to improve the clarity of the fitted functions. The fitted functions correspond to the parameters shown for Fit 1 in Table I and are plotted over only those data points that were included in the fitting procedure. The breakpoint of each fitted function (point of intersection of the two straight lines) is shown in parentheses after each entry in the figure legend.

and sixth columns of Table I (Fit 1). (The 3-dB-down cutoff for each straight-line fit is shown in the fifth column, but the usefulness of this quantity is limited when comparing across functions with different roll-off rates.) The best-fitting pairs of straight lines corresponding to these parameters are plotted among the data in Fig. 4. The plotted lines extend over only the data points that were included in the fitting procedure.

Recognizing that the estimated breakpoints and roll-offs will be influenced by the number of data points included in the fitting procedure, the straight-line fits were repeated with one less data point included for each TMTF (excluding the data point for the highest modulation frequency included in Fit 1). The breakpoints, 3-dB-down cutoffs, roll-offs, and percentages of variance accounted for are shown in the last three columns of Table I (Fit 2). Fits 1 and 2 provide a range of breakpoints for each TMTF and a more conservative basis for assessing differences across TMTFs than would a single estimate of the breakpoint.

As shown in Table I, the ranges of breakpoints across Fits 1 and 2 for the two monaural TMTFs are quite comparable, failing to overlap by only a few hertz. Table I shows that the roll-offs estimated for the monaural TMTF measured here are slightly steeper than those estimated for the TMTF of Kohlrausch *et al.* (2000). The shallower slope of the Kohlrausch *et al.* (2000) TMTF appears to be driven by the two points of the TMTF for the highest modulation frequencies included in the fitting procedure (see Fig. 4).

Comparison of the monaural and binaural TMTFs is complicated by the fact that there appears to be an effect of carrier bandwidth on the form of the interaural TMTF. As shown in Table I, the three interaural TMTFs span ranges of breakpoints across Fits 1 and 2 that differ greatly. For example, the midpoints of the breakpoint ranges for the narrowband noises differ by more than an octave. Furthermore, the roll-offs of the interaural TMTFs decrease with increasing carrier bandwidth. There is presently no obvious explanation for the effect of bandwidth on the interaural TMTF.

The interaural TMTFs for stimuli of different bandwidth are sufficiently different that it is probably inappropriate to average across them. Therefore, it seems that the most reasonable comparison that can be made between the monaural and interaural TMTFs is for those measured with a tonal carrier. As noted by Grantham (1984), the vertical separation between TMTFs reflects differences in overall sensitivity in the monaural and interaural tasks while differences in cutoff frequency represent differences in temporal resolution. In the present work, the ranges of breakpoints for the monaural and interaural tonal TMTFs were 116–121 and 79–85 Hz, respectively, indicating lower temporal resolution for the binaural system. (The breakpoints estimated here are compared to cutoff frequencies derived from previously published data in Sec. IV B.) This outcome is qualitatively consistent with the physiological observation that cells in the lateral superior olive that encode interaural level differences have a temporal limitation with a lower cutoff frequency than observed monaurally in their afferent fibers (Joris, 1996). Although these physiological cutoff frequencies are higher than those measured psychophysically, it seems plausible that this relative difference between monaural and binaural temporal resolution is maintained at higher neural levels of the auditory system where physiological measures of temporal resolution more closely match psychophysical performance.

While the monaural tonal TMTF rolls off continuously up to the f_m at which sideband detection provides a strong cue, the interaural tonal TMTF (shaded symbols, upper left panel of Fig. 3) rolls off between $f_m=70$ and 200 Hz then becomes relatively flat for $f_m=200$ –400 Hz, above which thresholds increase to the point of being unmeasurable. (Note that in the interaural conditions, in contrast to the monaural conditions, the mere presence of the modulation sidebands cannot serve as a cue for solving the listening task because the sidebands are present in all listening intervals.) The interaural TMTFs for diotic noise carriers (shaded symbols, middle and bottom left panels of Fig. 3) also do not roll-off monotonically above the estimated breakpoint. In fact, thresholds in the average interaural TMTF for the 300-Hz diotic noise carrier (shaded symbols, bottom left panel of Fig. 3) decrease dramatically above $f_m=200$ Hz. As indicated earlier, this is due primarily to the data of one listener. All three listeners' interaural TMTFs for the 300-Hz diotic noise are shown in Fig. 5. While the TMTFs for S1 and S2 show relatively little systematic variation across f_m , each varying over a range of less than 10 dB, the TMTF for S3 shows steeply improving (decreasing) thresholds at higher f_m 's.

One explanation that was considered in attempting to account for S3's performance at high frequencies is related to off-frequency listening. Perhaps S3 made use of the output of a critical-band filter that effectively equalized the levels of the carrier band and modulation sideband such that the modulation depth at the output of that filter would be greater than the nominal modulation depth of the stimulus (Goldstein, 1967). In order to pursue this possibility, S3 performed the binaural modulation detection task again for the 300-Hz diotic carrier at $f_m=650$ Hz, except with either the upper or lower modulation sideband filtered out of the stimulus.

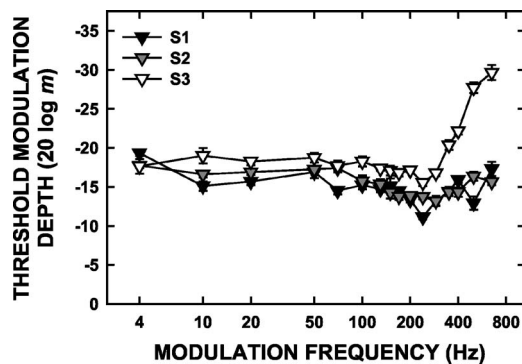


FIG. 5. The interaural TMTFs measured for the individual listeners for a diotic 300-Hz-wide band noise carrier. Error bars represent standard errors of the mean computed across four threshold estimates for each listener.

Thresholds were immeasurably large in these conditions. Whatever cue S3 made use of at high f_m with the 300-Hz diotic carrier apparently depended on the presence of both the upper and lower modulation sidebands.

Previous estimates of monaural time constants and binaural time constants (described in Sec. I) also indicate lower temporal resolution for binaural processing relative to monaural but the respective time constants correspond to cutoff frequencies that are much lower than the cutoff frequencies (breakpoints) estimated here. To facilitate comparison to the present results, the time constants (τ) estimated by Grantham and Wightman (1979), Kollmeier and Gilkey (1990), and Culling and Summerfield (1998) were converted to 3-dB-down cutoff frequencies (f_c) assuming a first-order, lowpass filter so that $f_c = 1/(2\pi\tau)$. Grantham and Wightman (1979) estimated τ between 44 and 243 ms, corresponding to values of f_c between 0.66 and 3.62 Hz. For the results of Kollmeier and Gilkey (1990), the corresponding range of binaural f_c is 3.73–9.88 Hz, and the range of monaural f_c is 10.83–26.98 Hz. A simple conversion of the time constants reported by Culling and Summerfield (1998) for the two sides of an asymmetric temporal window yields a range of f_c from 2.6 to 13.6 Hz. It is possible that the large discrepancy between cutoff frequencies estimated for the data from the present study and from the previous studies indicates that the previous studies measured a fundamentally different aspect of binaural temporal resolution than the present experiment. If one considers an interaurally uncorrelated noise to have a time-varying instantaneous ILD (as depicted in Fig. 1), then variation in the value of interaural correlation is effectively modulation of the interaural modulator, or second-order interaural modulation. Lorenzi *et al.* (2001) compared first-order and second-order monaural TMTFs. When detecting first-order monaural modulation, listeners discriminate between an unmodulated carrier and one that is sinusoidally amplitude-modulated. Detection of second-order modulation involves discrimination between a stimulus that is sinusoidally amplitude-modulated at a fixed modulation depth and a stimulus for which the modulation depth varies over time. Lorenzi *et al.* (2001) showed that cutoff frequencies implied by the low-pass segments of second-order TMTFs were lower than those for first-order TMTFs. Perhaps in a similar way the binaural system exhibits better temporal resolution

for detection of static interaural modulation (measured here) than for time-varying interaural modulation (as measured in the studies cited above).

A difference in overall sensitivity in the monaural and interaural modulation-detection tasks was observed in the present experiment (as indicated by the vertical separation between the TMTFs of Fig. 4). In contrast, Stellmack *et al.* (2004) reported that thresholds for binaural discrimination of static intensity in a single-interval task were nearly equal to monaural intensity-discrimination thresholds measured in a two-interval task. This further supports the idea that the monaural and interaural TMTFs represent the temporal characteristics of two different processes.

Although the sources of all the differences between the monaural and interaural TMTFs are not apparent, a clear result is that the temporal resolution characteristics of the binaural system for detection of fluctuating interaural intensity cannot be extrapolated directly from the pure-tone monaural TMTF. For the TMTFs for which direct comparison seems most appropriate, the breakpoint of the interaural tonal TMTF is substantially lower than the monaural tonal TMTF. In addition, there are differences between the interaural TMTFs across carrier bandwidths that cannot be explained at this time.

B. Influence of intrinsic fluctuations on the form of the TMTFs

As noted earlier, the monaural TMTFs measured with noise carriers differed from the tonal TMTF (see the upper right panel of Fig. 3) in ways that can be explained on the basis of intrinsic fluctuations in the carrier masking or interfering with the signal modulation. Plotted in the top panel of Fig. 6 are the differences between monaural thresholds measured with the tonal carrier and those measured with the 30-Hz-wide noise (diamonds) and 300-Hz-wide noise (triangles), that is, the amount of masking in each case. In the top panel of Fig. 7, average monaural envelope power spectra for 1000 independent samples of narrowband Gaussian noise (each 1 s in duration and windowed with 150-ms raised-cosine on-off ramps) are shown for bandwidths of 30 and 300 Hz, with the total power equal at the two bandwidths. It can be seen that energy in the envelope spectrum is concentrated below the frequency corresponding to the bandwidth of the noise (which also can be demonstrated analytically, e.g., Lawson and Uhlenbeck, 1950). (The specific shape of the spectra in the top panel of Fig. 7 below about 5 Hz depends upon the form and duration of the onset-offset ramps of the stimuli.) The crossover of the two functions in the top panel of Fig. 6 occurs at approximately the modulation frequency at which the envelope power spectra cross in the top panel of Fig. 7. The intrinsic fluctuations in the noise envelope presumably interfere with or mask the signal modulation to the largest extent when the signal f_m is less than the bandwidth of the carrier (Dau *et al.*, 1999). As described in Sec. I, an additional assumption of some type of limited modulation-frequency selectivity must be made in order to account for the form of the TMTFs on the basis of masking by the intrinsic fluctuations of the noise carrier, re-

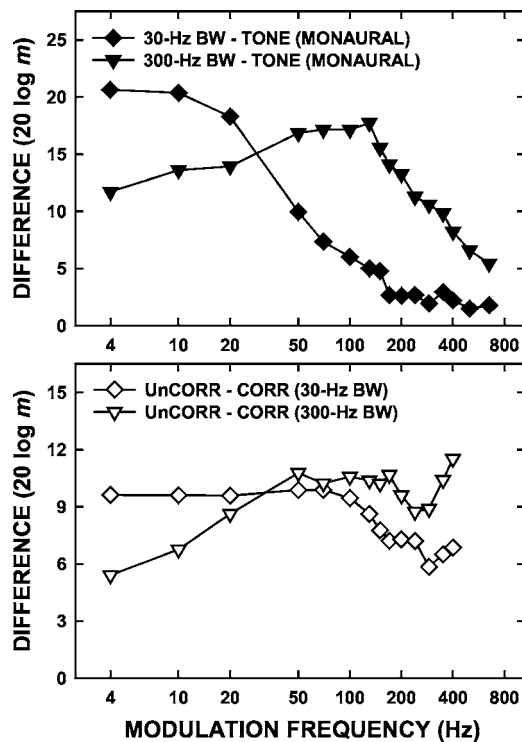


FIG. 6. Top panel: Difference between monaural modulation-detection thresholds measured with a narrowband-noise carrier (diamonds, 30-Hz wide; triangles, 300-Hz wide) and those measured with a tonal carrier. Bottom panel: Difference between interaural modulation-detection thresholds measured with interaurally uncorrelated noise carriers and interaurally correlated noise carriers (diamonds, 30-Hz wide; triangles 300-Hz wide). All differences were computed on the mean data shown in Fig. 3.

sulting in more masking when masker energy is present in the modulation spectrum near the modulation frequency of the signal.

In order to assess the effects of intrinsic interaural fluctuations of the carrier on the interaural TMTF for uncorrelated noise carriers, the differences between thresholds measured with correlated and uncorrelated 30- and 300-Hz-wide noise carriers were computed and plotted in the bottom panel of Fig. 6. The amount of masking was larger at frequencies below $f_m=50$ Hz for the 30-Hz-wide noise than for the 300-Hz-wide noise, as was the case for the monaural TMTFs (see the top panel of Fig. 6), although the overall amount of masking in the interaural conditions was less than in the monaural conditions. The curves in the bottom panel of Fig. 7 represent the average spectra of the instantaneous ILD functions for carrier bandwidths of 30 and 300 Hz. The spectra were computed by generating two independent narrowband Gaussian noise samples (each 1 s in duration, windowed with 150-ms raised-cosine on-off ramps), representing uncorrelated left- and right-ear noise carriers. The instantaneous ILD function was computed from the Hilbert envelopes of the left- and right-ear carriers as described by Eq. (1) and the spectrum of that function then was computed via FFT. Spectra were computed in this way and averaged for 1000 randomly generated pairs of noise yielding the mean spectra plotted in the bottom panel of Fig. 7. The bottom panel of Fig. 7 shows that, as was the case for intrinsic fluctuations in monaural intensity, energy in the spectra of

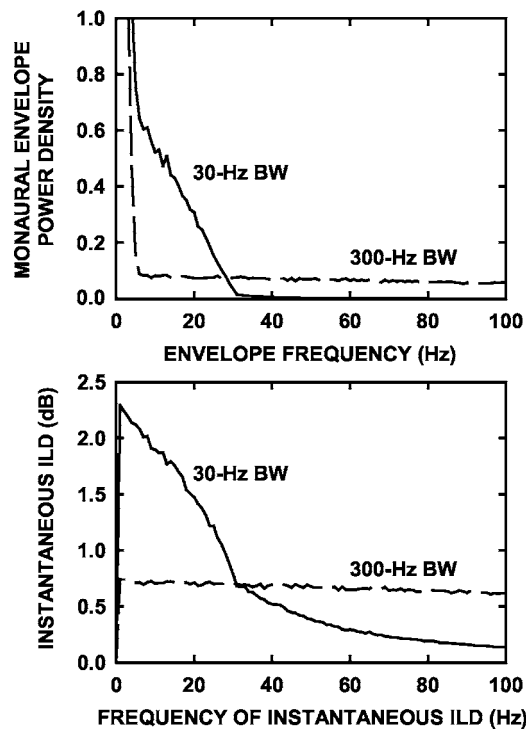


FIG. 7. Top panel: Monaural envelope power spectra averaged over 1000 random noise samples for 30-Hz-wide noise (solid line) and 300-Hz-wide noise (dashed line), where the total power was equal at the two bandwidths. Bottom panel: Spectra of instantaneous ILD functions averaged over 1000 samples of interaurally uncorrelated noise with bandwidths of 30 Hz (solid line) and 300 Hz (dashed line). Instantaneous ILD functions were computed from the Hilbert envelopes of the left- and right-ear wave forms using Eq. (1). In all cases, the stimuli were 1 s in duration and were windowed with 150-ms raised-cosine on-off ramps. Spectra were obtained by computing FFTs on digital waveforms sampled at 44.1 kHz.

the intrinsic interaural fluctuations is concentrated at frequencies below the bandwidth of the carrier. For both the monaural and interaural TMTFs, the amount of masking as a function of f_m (Fig. 6) appears to be related to the corresponding spectrum of the intrinsic carrier fluctuations (Fig. 7).

Comparison of the tonal and diotic noise interaural TMTFs (shaded symbols, lower right panel of Fig. 3) seems to reveal little effect of intrinsic monaural fluctuations in and of themselves on the interaural TMTF for f_m up to about 200 Hz. Thresholds are slightly elevated for the lower f_m 's for the diotic noise carriers relative to those for the tonal carrier, but this small difference does not appear to depend on the frequency range of the intrinsic monaural fluctuations in the carrier, as shown in the top panel of Fig. 7. For example, with the 30-Hz-wide correlated carrier, the effect of intrinsic monaural fluctuations on interaural thresholds at low f_m where substantial energy exists in the spectrum of the carrier envelope (shaded diamonds versus shaded circles in the lower right panel of Fig. 3) is much smaller than that in the corresponding monaural TMTF (diamonds versus squares in the upper right panel of Fig. 3).

These results suggest that intrinsic monaural or binaural fluctuations due to a noise carrier exhibit the most masking when those fluctuations occur in the same domain (monaural or binaural) as the sinusoidal signal modulation. For the

monaural TMTFs (top right panel of Fig. 3), thresholds increase by the largest amount for the noise carriers (relative to those for the tonal carrier) over the range of f_m in which there are substantial intrinsic fluctuations in the carrier. For the interaural TMTFs measured with uncorrelated noise carriers (open symbols, bottom right panel of Fig. 3), the largest change in thresholds (relative to those for the diotic carriers), or the most masking, occurs over the range of f_m in which there are intrinsic fluctuations in interaural level, although the degree of “tuning” (represented by the variation in the amount of masking with changes in f_m) is much smaller than for the monaural TMTFs. In particular, for the interaural TMTF for the 30-Hz-wide noise, the amount of masking remains nearly constant for modulation frequencies up to about 70 Hz. The differences between the amounts of masking in the monaural and binaural conditions may be due in part to the fact that for the 30-Hz bandwidth, the spectrum of the instantaneous ILD function (bottom panel of Fig. 7) rolls off more gradually than the monaural envelope power spectrum (top panel of Fig. 7) and nonzero ILDs are present at frequencies well above the nominal bandwidth of the carrier. In addition, the fact that the amount of masking remains high well above the f_m corresponding to the carrier bandwidth may indicate that any frequency-selective mechanism in the binaural domain may be more broadly tuned than that for monaural modulation.

Grantham and Bacon (1991) performed a study of “Binaural modulation masking” which is somewhat related to the present experiment, but it does not address the issue of masking produced by intrinsic interaural fluctuations in the same manner as the present experiment. They first measured thresholds for detection of monaural signal modulation in the presence of a monaural masker modulator, both modulating a common broadband noise carrier. The masker f_m was fixed at 16 Hz and the signal f_m ranged from 2 to 512 Hz. They then measured thresholds for detection of interaurally out-of-phase modulation (relative to no signal modulator) in the presence of a diotic masker modulator for an interaurally uncorrelated broadband noise carrier. These monaural and binaural conditions would correspond to the binaural masking-level difference conditions $N_m S_m$ and $N_0 S_\pi$, respectively, in the modulation domain. Both sets of thresholds showed tuning in that the greatest amount of masking occurred when the signal and masker modulation frequencies were equal. In addition, a binaural MLD was observed in that the mean $N_0 S_\pi$ thresholds were consistently lower than the $N_m S_m$ thresholds. While the data of Grantham and Bacon (1991) show that thresholds for detection of dichotic modulation depend upon the relative modulation frequencies of the signal and masker modulation, it is difficult to relate their $N_0 S_\pi$ thresholds to any of the conditions run in the present experiment. The $N_0 S_\pi$ thresholds of Grantham and Bacon (1991) involved detection of a dichotic modulator while the binaural conditions of the present experiment involved discrimination of interaural phase of the modulator. In other words, in the present binaural conditions, there were no potential monaural cues available which is not true in the Grantham and Bacon (1991) experiment.

C. Role of center frequency in the measurement of the interaural TMTF

Grantham (1984) measured TMTFs primarily with uncorrelated noise carriers for low- and high-frequency carriers of varying bandwidths. However, Grantham (1984, footnote 2) did describe pilot data comparing thresholds gathered with correlated and uncorrelated noise carriers centered at 1 kHz. The forms of the interaural TMTFs measured here for the diotic and uncorrelated noise carriers seem to differ from those described by Grantham. Grantham states that the TMTF for a diotic noise carrier centered at 1 kHz showed a sharp upturn (decrease in thresholds) for $f_m > 50$ Hz. He concluded that this was due to detection of the modulation on the basis of interaural decorrelation of the stimulus for high f_m . In contrast, the interaural TMTFs measured with diotic carriers centered at 5 kHz in the present experiment (shaded symbols of Fig. 3) were nearly flat (within a few decibels) for f_m up to about 70 Hz, above which thresholds tended to increase. Grantham (1984) also stated that there were little or no differences between the TMTFs for diotic and uncorrelated noise carriers centered at 1 kHz for $f_m < 50$ Hz, while in the present experiment the TMTFs for diotic and uncorrelated noise carriers centered at 5 kHz differed by more than 10 dB over the same range of f_m .

The different center frequencies used by Grantham (1984) to gather the data described above (from Grantham’s footnote 2) and in the present experiment for a diotic noise carrier may be important. Grantham’s diotic carrier centered at 1 kHz was over 250 Hz wide (at its 3-dB bandwidth). Given that the equivalent rectangular bandwidth of the auditory filter is about 130 Hz at 1 kHz (Glasberg and Moore, 1990), this means that Grantham’s modulated stimuli spanned several critical bands and, perhaps more important, the modulation sidebands fell in a frequency region in which listeners are sensitive to interaural phase differences in the fine structure. For the results described by Grantham for a diotic carrier, regardless of the modulation frequency, listeners may have responded to interaural phase differences in the fine structure of the modulation sidebands that are produced when the phase of the signal modulator was interaurally changed. Such cues were not available in the high-frequency stimuli of the present experiment. This may account for the sharply decreasing thresholds for $f_m > 50$ Hz reported by Grantham, while in most cases in the present data thresholds increased with increasing modulation frequency. However, Grantham (1984) also reported almost no difference between thresholds measured with correlated and uncorrelated carriers for $f_m < 50$ Hz, while in the present experiment thresholds were up to 10 dB lower for the correlated carriers relative to the uncorrelated carriers. The explanation put forth above regarding sensitivity to interaural phase differences in the fine structure of the sidebands would seem to predict the opposite result: better diotic thresholds for the low-frequency carriers of Grantham and no difference between diotic and dichotic thresholds for the high-frequency carriers of this experiment. Thus, although it is true that fine-structure cues exist in the modulated low-frequency carriers but not the

high-frequency carriers, this fact does not appear to account for all differences between the data gathered with low- and high-frequency carriers.

V. CONCLUSIONS

- (1) Temporal resolution for fluctuating interaural intensity differences was poorer than monaural temporal resolution for intensity fluctuations. For the present data, the breakpoint of the interaural TMTF appeared to be about a half-octave lower than the breakpoint of the monaural TMTF for a 5-kHz pure-tone carrier. Breakpoints were estimated for two subsets of the interaural and monaural data and ranged from 79 to 85 Hz for the interaural TMTF and from 116–121 Hz for the monaural TMTF. Roll-off rates of the fitted functions were approximately equal for the interaural and monaural TMTFs.
- (2) The breakpoints and roll-off rates of the interaural TMTFs measured with interaurally correlated narrow-band noise carriers differed from those for the interaural TMTF measured with a tonal carrier. The source of these differences is not clear.
- (3) Consistent with the results of previous research, the monaural TMTFs measured with a noise carrier exhibited higher thresholds (relative to monaural TMTFs measured with a pure-tone carrier) at modulation frequencies at which there are intrinsic fluctuations in the envelope of the carrier. In a similar way, the interaural TMTFs measured with uncorrelated noise carriers showed effects consistent with interference produced by intrinsic interaural fluctuations in the noise carrier. Specifically, thresholds were slightly elevated for f_m 's at which there are larger intrinsic fluctuations in the interaural intensity differences of the carrier. The masking functions (amount of masking produced by intrinsic fluctuations as a function of signal modulation frequency) for the monaural and interaural TMTFs differed in form. Intrinsic monaural fluctuations generally produced more masking than binaural fluctuations at lower signal modulation frequencies.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Armin Kohlrausch, Dr. Wesley Grantham, and an anonymous reviewer for their helpful comments. This work was supported by Research Grant No. R03 DC 05343-01 and Research Grant No. R01 DC 00683, both from the National Institute on Deafness and Communication Disorders, National Institutes of Health.

¹The choice of which points of each TMTF would be included in each fit was arbitrary to some extent. The stated criterion for data points to be included was to include those points up to the f_m at which the slope of the TMTF changes from positive to negative. Strictly speaking, the slope of the monaural and interaural TMTFs for the tone changed from positive to negative at $f_m=20$ Hz, while the slopes of the two correlated noise interaural TMTFs changed sign at $f_m=10$ Hz. In practice, the slope-change criterion was applied to each TMTF above the f_m at which each function

appeared to show a systematic rolloff or clear lowpass characteristic, determined by eye. The fitted functions are plotted through only those points used to obtain the fits in Fig. 4.

- 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.
- Bernstein, L. R., and Trahiotis, C. (2005). "Processing of interaural temporal disparities with both 'transposed' and conventional stimuli," in *Auditory Signal Processing: Physiology, Psychoacoustics, and Models* edited by D. Pressnitzer, A. de Cheveigne, S. McAdams, and L. Collet (Springer, New York, in press).
- Culling, J. F., and Summerfield, Q. (1998). "Measurements of the binaural temporal window using a detection task," *J. Acoust. Soc. Am.* **103**, 3540–3553.
- Dau, T., Kollmeier, B., and Kohlrausch, A. (1997). "Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers," *J. Acoust. Soc. Am.* **102**, 2892–2905.
- Dau, T., Verhey, J., and Kohlrausch, A. (1999). "Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers," *J. Acoust. Soc. Am.* **106**, 2752–2760.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Goldstein, J. L. (1967). "Auditory spectral filtering and monaural phase perception," *J. Acoust. Soc. Am.* **41**, 458–479.
- 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. (1979). "Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation," *J. Acoust. Soc. Am.* **65**, 1509–1517.
- Joris, P. X. (1996). "Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive," *J. Neurophysiol.* **76**, 2137–2156.
- Klump, 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.
- Kollmeier, B., and Gilkey, R. H. (1990). "Binaural forward and backward masking: Evidence for sluggishness in binaural detection," *J. Acoust. Soc. Am.* **87**, 1709–1719.
- 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 psychophysics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lorenzi, C., Soares, C., and Vonner, T. (2001). "Second-order temporal modulation transfer functions," *J. Acoust. Soc. Am.* **110**, 1030–1038.
- Nuetzel, J. M., and Hafter, E. R. (1976). "Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration," *J. Acoust. Soc. Am.* **60**, 1339–1346.
- Rickert, M. E., and Viemeister, N. F. (1998). "Temporal versus spectral cues in AM detection," *J. Acoust. Soc. Am.* **103**, 2842.
- Stellmack, M. A., Viemeister, N. F., and Byrne, A. J. (2004). "Monaural and interaural intensity discrimination: Level effects and the 'binaural advantage'," *J. Acoust. Soc. Am.* **116**, 1149–1159.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.
- Viemeister, N. F., Rickert, M., Law, M., and Stellmack, M. (2002). "Psychophysical and physiological aspects of auditory temporal processing," in *Genetics and the Function of the Auditory System*, Proceedings of the 19th Danavox Symposium, edited by L. Tranebjaerg, J. Christensen-Dalsgaard, T. Andersen, and T. Poulsen (Holmens Trykkeri, Denmark), pp. 273–291.
- Wiegand, L., and Patterson, R. D. (1999). "Quantifying the distortion products generated by amplitude-modulated noise," *J. Acoust. Soc. Am.* **106**, 2709–2718.
- Yin, T. C. T. (2002). "Neural mechanisms of encoding binaural localization cues in the auditory brainstem," in *Integrative Functions in the Mammalian Auditory Pathway*, edited by D. Oertel, R. R. Fay, and A. N. Popper (Springer, New York), pp. 99–159.