Interaural Correlation Fails to Account for Detection in a Classic Binaural Task: Dynamic ITDs Dominate N0Sπ Detection

Marcel van der Heijden[1,2] and Philip X. Joris[1]

[1]Laboratory of Auditory Neurophysiology, K.U. Leuven Medical School, Leuven, 3000, Belgium
[2]Department of Neuroscience, Erasmus MC, University Medical Center Rotterdam, 3000 CA, Rotterdam, Netherlands

ABSTRACT

Binaural signal detection in an NoSπ task relies on interaural disparities introduced by adding an antiphasic signal to diotic noise. What metric of interaural disparity best predicts performance? Some models use interaural correlation; others differentiate between dynamic interaural time differences (ITDs) and interaural level differences (ILDs) of the effective stimulus. To examine the relative contributions of ITDs and ILDs in binaural detection, we developed a novel signal processing technique that selectively degrades different aspects (potential cues) of binaural stimuli (e.g., only ITDs are scrambled). Degrading a particular cue will affect performance only if that cue is relevant to the binaural processing underlying detection. This selective scrambling technique was applied to the stimuli of a classic N0Sπ task in which the listener had to detect an antiphasic 500-Hz signal in the presence of a diotic wideband noise masker. Data obtained from five listeners showed that (1) selective scrambling of ILDs had little effect on binaural detection, (2) selective scrambling of ITDs significantly degraded detection, and (3) combined scrambling of ILDs and ITDs had the same effect as exclusive scrambling of ITDs. Regarding the question which stimulus properties determine detection, we conclude that for this binaural task (1) dynamic ITDs dominate detection performance, (2) ILDs are largely irrelevant, and (3) interaural correlation of the stimulus is a poor predictor of detection. Two simple stimulus-based models that each reproduce all binaural aspects of the data quite well are described: (1) a single-parameter detection model using ITD variance as detection criterion and (2) a compressive transformation followed by a crosscorrelation analysis. The success of both of these contrasting models shows that our data alone cannot reveal the mechanisms underlying the dominance of ITD cues. The physiological implications of our findings are discussed.

Keywords: binaural detection, masking, ITD, ILD, MLD, binaural modulation

Correspondence to: Marcel van der Heijden · Department of Neuroscience, Erasmus MC · University Medical Center Rotterdam · 3000 CA, Rotterdam, Netherlands. Telephone: +31-10-7043309; fax: +31-10-7044734; email: m.vanderheyden@erasmusmc.nl

INTRODUCTION

The classic example of binaural detection is an N0Sπ task, in which an interaurally phase-reversed tone Sπ is masked by a diotic (interaurally identical) noise N0 (Hirsh 1948). When the masker is a wideband noise and the signal a low-frequency (<1 kHz) tone, the binaural advantage re the diotic reference condition, N0S0, is 10–15 dB. This binaural advantage is called the binaural masking level difference (MLD). It reflects the listeners' sensitivity to the interaural disparities caused by adding the Sπ signal to the diotic masker. Apparently, the brain performs a real-time comparison of the stimuli received by the two ears and is sensitive to tiny deviations from perfect interaural correlation. This ability is relevant in everyday hearing conditions because the wavelength of low-frequency sounds is large compared to the distance between the ears, causing a high degree of "baseline" interaural correlation regardless of the azimuthal

angle of incidence. If the hearing system is to benefit from having two ears, the ability to detect tiny deviations from interaural equality at low frequencies is a necessity. Therefore, despite the somewhat contrived appearance of the N0Sπ stimulus configuration, it captures and quantifies a fundamental ability of the binaural system.

Determining MLDs, however, does not shed much light on the mechanisms that underlie binaural signal detection, nor does it reveal which aspects of the stimulus ("cues") are important for detection and which are not. Any quantitative model or theory of binaural detection must deal with the basic question of how to quantify interaural disparities of auditory stimuli (Colburn and Durlach 1978). In the classic stimulus-oriented theories of binaural detection, two alternative approaches prevail. In the first approach, the disparities are quantified in terms of normalized interaural correlation of the (effective) stimulus, and detection is associated with a critical change in correlation caused by the signal (Cherry and Sayers 1959; Dolan and Robinson 1967; Osman 1971).

The second approach distinguishes two types of binaural disparity: interaural time differences (ITDs) and interaural level differences (ILDs). With this approach, it is taken into account that only a limited band of noise centered around the signal frequency is effective in masking the tonal signal. This band of noise can be represented by a tone with slowly varying amplitude and phase. For a diotic (N0) noise, these amplitude and phase modulations are identical in the two ears, but the addition of an Sπ signal to the N0 noise introduces slowly varying, random, ITDs and ILDs. (In the present context of detection of a narrowband signal, the distinction of interaural phase and interaural time is unimportant.) It is the listeners' sensitivity to these dynamic interaural disparities that is assumed to underlie binaural detection tasks in the second type of models. Webster (1951) hypothesized that binaural detection of tones in broadband noise was based on the detection of a certain threshold value for detecting ITDs. Subsequent work (Jeffress et al. 1956) resulted in an estimate of 100 μs for this threshold ITD. Later extensions also included the processing of dynamic ILDs (Durlach 1964; Hafter and Carrier 1970). Models of this are generally refferred to as "lateralization models of binaural detection" because they attempt to connect binaural signal detection with sound localization. In many models of binaural detection that are inspired by physiological data (e.g., Colburn 1973), the different roles of ILDs and ITDs are not described explicitly, but implicitly in the form of assumptions concerning peripheral processing, sources of internal noise, and decision criteria.

Another way of introducing an effective asymmetry between ITDs and ILDs is by postulating a stage of

peripheral envelope compression prior to the binaural interaction (Van de Par and Kohlrausch 1998; Bernstein et al. 1999). Envelope compression reduces the size of the amplitude fluctuations in each monaural channel without affecting the phase of the waveforms, thereby causing a reduction in the magnitude of ILDs at the input of the binaural processor and leaving ITDs unaffected. If the binaural processor itself is equivalent to a crosscorrelator (which does not discriminate between ILDs and ITDs), peripheral envelope compression will cause an effective dominance of ITDs.

In view of these different theories and models, several of which postulate an asymmetry between ITDs and ILDs, it is imperative to find out whether such an asymmetry actually exists. This raises the question of how the relative importance of dynamic ITDs and ILDs in binaural detection can be empirically assessed. Unfortunately, ITDs and ILDs are inextricably linked in most common binaural detection tasks. In the case of Sπ tones masked by diotic Gaussian noise, the overall magnitudes of both ITDs and ILDs grow with signal level. This covariation of ILDs and ITDs (Fig. 1A) leads to an effective equivalence between correlation-based models, in which ILDs and ITDs merge into the single correlation metric, and models that do discriminate between dynamic ITDs and ILDs. Both types of models predict the same dependence of binaural thresholds on the basic set of stimulus parameters (Domnitz and Colburn 1976). Of course, this equivalence of predicted thresholds does not in any way reduce the importance of the ITD-versus-ILD issue itself, which is a fundamental question of binaural mechanisms transcending the
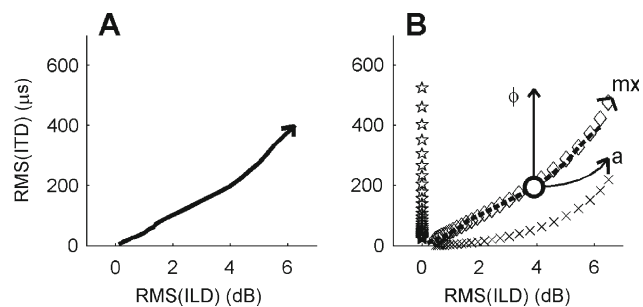
FIG. 1. The simultaneous growth of the magnitudes of dynamic ITDs and ILDs evoked by adding an antiphasic tone Sπ to a diotic noise N0 (**A**). The *arrowhead* indicates the direction of increasing Sπ/N0 ratio. This curve provides the gauge for comparing the relative strengths of ITDs and ILDs (see text). **B** The combined effect of binaural modulation and the addition of the Sπ tone. *Different symbols* indicate different modulation types: binaural QFM (*stars*), binaural AM (*x marks*), and mixed binaural modulation (*diamonds*). The *dashed line* is replotted from **A**. The *white circle* indicates an N0Sπ stimulus with S/N=−13 dB; the *arrows* show how ITDs and ILDs are affected when this N0Sπ stimulus is subjected to different types of binaural modulation. In all cases, the binaural metrics were evaluated on a 100-Hz-wide noise band around 500 Hz. The magnitudes are expressed as RMS over the stimulus duration.

domain of MLDs. The equivalence does, however, present a stumbling block in designing empirical tests of the relative roles of ITDs and ILDs.

One way to disentangle the roles of ITDs and ILDs is the use of stimulus configurations, which, contrary to the customary tone-in-Gaussian-noise configurations, allow the separate control of ITDs and ILDs. McFadden et al. (1971) reported binaural detection of a narrowband noise signal in the presence of a narrowband noise masker. The antiphasic signal was a binaurally phase-shifted version of the masker, and by varying the phase difference between signal and masker, the relative contribution of ITDs and ILDs was manipulated. No systematic dominance of either cue was found. Note, however, that ITDs and ILDs occurring in this stimulus configuration are static and systematic. In contrast, in a typical binaural tone-in-noise task, the signal introduces dynamic, random, variations in ITDs and ILDs. The relevance of these results to classic, wideband, N0Sπ detection is, therefore, unclear.

Van de Par and Kohlrausch (1998) performed a dynamic version of the experiment of McFadden et al. They used multiplied noise (i.e., low-pass noise multiplied by a sinusoid) to mask tonal signals having a fixed phase re the masker carrier. Again, the phase angle determines the relative contribution of ITDs and ILDs, but this time, their variation is dynamic and random. At low frequencies, no dominance of either ITDs or ILDs was found, leading to the conclusion that interaural correlation was a good determinant of low-frequency binaural signal detection. Cochlear filtering, however, limits this type of tests to sub-critical bandwidths. Moreover, there are large differences in waveform statistics between multiplied and Gaussian noise (Van der Heijden and Kohlrausch 1995). Again it is unclear to what extent the conclusions of Van de Par and Kohlrausch (1998) may be generalized to a classic wideband N0Sπ condition.

The present study takes a different approach. Most binaural models postulate some kind of internal noise or "jitter" serving to limit performance. In the present study, the jitter is explicitly imposed on the *stimulus* with the purpose of assessing its disruptive effect on binaural detection. Starting with a classic wideband N0Sπ task, the stimuli are distorted (jittered) in different ways that selectively scramble the different cues: ITDs, ILDs, or both. The decline in performance caused by the different types of distortion then reveals the relative importance of the distorted cues. Specifically, the distortion consists of a mixed modulation of the complete stimulus waveform, and the types of distortion differ in the interaural phase of the modulation constituents. Our findings unambiguously point to a strong dominance of dynamic ITDs in the wideband masking condition considered.

METHODS

Quantifying dynamic ITDs and ILDs in an N0Sπ stimulus

One might wonder how the relevant importance of dynamic ILDs or ITDs in a binaural task can be assessed at all. ILDs and ITDs are apples and oranges: They are measured in different units, so their magnitudes cannot be directly compared. The real question, however, is their relative importance *in the context of a given binaural task*. The present study deals with an N0Sπ detection task, so the correct way of comparing the magnitudes of ILDs and ITDs is to consider how each of them varies with signal level. To quantify the magnitude of fluctuations of ILDs and ITDs, the root mean square (RMS) of each of these fluctuating quantities in a 100-Hz-wide band of noise around 500 Hz was determined. Both RMS values grow with increasing level of the Sπ signal (Zurek 1991). Figure 1A shows their covariation with signal level by plotting RMS(ITD) directly against RMS (ILD). As will be shown below, this covariation of ILDs and ITDs under the "natural conditions" of the N0Sπ task provides the baseline for assessing their manipulation by the binaural modulation.

Binaural modulation

The general idea of the binaural modulation is as follows. The stimulus to each ear is modulated in a manner that affects both the amplitude and the phase of the waveform. The amplitude modulation introduces maxima and minima in the envelope; the phase modulation introduces phase leads and phase lags re the unmodulated waveform. As long as the modulation is identical in the two ears, it does not introduce any binaural differences. Conversely, when the modulation is completely out of phase between the ears, the envelope maxima in one ear will coincide with envelope minima in the other ear; similarly, phase leads in one ear will coincide with lags in the other. Thus both envelopes and phase shifts are interaurally opposite, leading to dynamic ILDs and ITDs. Interestingly, it is also possible to choose modulation phases in the two ears in such a way that the envelopes are interaurally reversed, but the phases are identical: This produces ILDs but no ITDs. Yet another choice of modulation phases yields interaurally identical envelopes and reversed phase shifts. Together the different types of binaural modulation realize the specific scrambling of binaural cues that is exploited in this study.

The mixed modulation used in this study is illustrated for single tones with the vector diagrams in Figure 2 (see also "Appendix", Eqs. 1–4). In panels A and B, the thick vertical arrow represents a carrier at frequency f_c (Jeffress et al. 1956). The two smaller
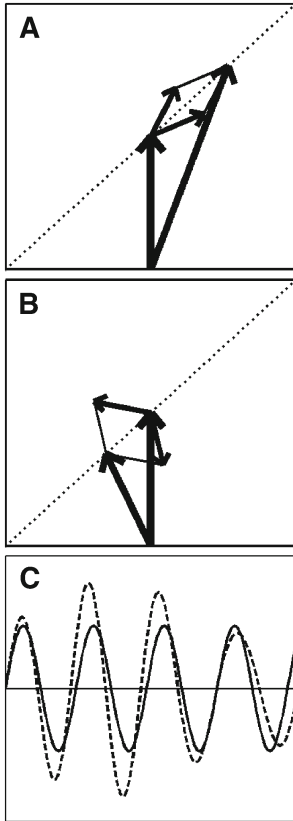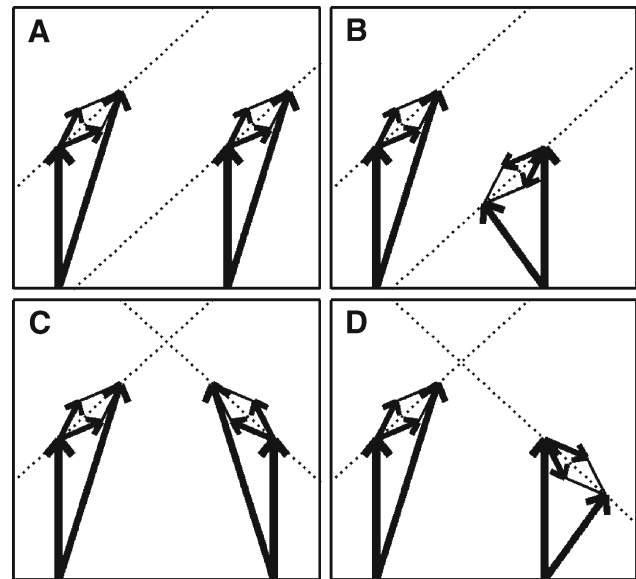
FIG. 2. Mixed modulation. **A**, **B** Vector diagrams. The *thick vertical arrow* represents the carrier. The *two small arrows* are the sidebands, which, due to their frequencies relative to the carrier, are spinning in opposite directions. **A** and **B** Snapshots at different time instants. The initial phases are chosen such that the resultant (i.e., the sum of carrier and sidebands, indicated by the *large skewed arrow*) is confined to the diagonal indicated by the *dotted line*. **C** Waveforms of the unmodulated carrier (*solid line*) and the modulated carrier (*dashed line*). Note that both the amplitude and the phase (zero crossings) are affected by the modulation.

arrows represent the sidebands at $f_c \pm f_m$, where f_m is the modulation frequency. Their angles represent their starting phases re the carrier. Their opposite rotations re the carrier confine the resultant vector (i.e., carrier plus sidebands; large skewed arrow) to the dotted diagonal line. Panels A and B show different time instances (snap shots) of the same modulation cycle. The modulation is *mixed* because, in the course of a modulation cycle, both amplitude and phase of the resultant are modulated with respect to the unmodulated carrier. This is clearly visible in the waveforms shown in panel C. An alternative description of this type of mixed modulation is provided by the decomposition of the sidebands in an amplitude-modulating (AM) pair (starting parallel to the carrier) and a quasi-frequency modulating (QFM) pair (starting perpendicular to the carrier). Their sum, having a 45° orientation re the carrier, represents an "equal mix" of AM and QFM.

The panels of Figure 3 show the different interaural combinations of mixed modulation used in this

study. The modulation was either identical in the two ears (panel A), completely opposite in the two ears (panel B), or different by 90°, in which case either the QFM component (panel C) or the AM component (panel D) is interaurally phase-reversed, while the other component is in phase. Note that in all four configurations depicted, the stimulus to each ear contains a mixed modulation. It is only in their *binaural* relations that the configurations express their essential differences. Based on this binaural aspect of the modulation, the configurations are named: *diotic* modulation (panel A), *mixed* binaural modulation (panel B), binaural *phase* modulation (panel C), and binaural *amplitude* modulation (panel D). Throughout this study, the modulation types are indicated with subscripts d, mx, ϕ, and a, respectively.

For large values of modulation depth, when the amplitudes of the sideband approach that of the carrier, the binaural amplitude modulation of Figure 3D also introduces small interaural phase differences, as is illustrated by the interaural inequality of the angles between carrier and resultant in Figure 3D. The spurious interaural phase modulation is a second-order effect, which only becomes important at large modulation depths. (This spurious phase modulation is analogous to the small amount of amplitude modulation that occurs when applying QFM; like the latter effect, the rate of the spurious modulation is twice the

FIG. 3. Binaural modulation. Each panel shows a pair of vector diagrams representing the stimuli presented to the two ears. The stimulus to each ear contains a mixed modulation as in Figure 2. Different choices of interaural phases of the sidebands lead to different binaural modulations. **A** Diotic modulation: modulations are interaurally identical; **B** mixed binaural modulation: both phase and amplitude modulation are interaurally phase-reversed; **C** binaural QFM: phase modulation is interaurally phase-reversed, while amplitude modulation is diotic; **D** binaural AM: amplitude modulation is phase-reversed, while phase modulation is interaurally in phase (though not completely identical).

rate of the proper modulation). This artifact could have been avoided by applying pure, antiphasic, QFM in both ears, but that strategy would introduce monaural differences between the stimulus conditions that would complicate the interpretation of the data. The potential effect of the spurious interaural phase modulations will be analyzed later.

Using binaural modulation to "scramble" an N0Sπ stimulus

So far, modulation of single tones was described. In the experiment, however, the different types of modulation were imposed on the complete waveforms comprising the stimuli of an N0Sπ (or N0S0) task. This was realized by using complex-analytic versions of the stimuli (see "Appendix", Eqs. 11a and 11b). Importantly, both masker and signal (if present) were simultaneously subjected to the same modulation. Such comodulation of signal and masker ensures that their mutual interaction is not altered by the modulation. This is crucial: For instance, had we modulated the noise but not the signal, then the signal would have become more audible during the envelope minima of the masker ("listening in the valleys"; Buus 1985).

Apart from the modulation, which is utilized here as a *post hoc* manipulation of the stimuli, the experiments involved conventional N0Sπ and N0S0 detection tasks, the details of which are specified below. The following notation is used for the different stimulus configurations. The basic (unmodulated) condition was either N0Sπ or N0S0. The modulation type is indicated by a subscript, e.g., $(N0S\pi)_d$ denotes diotic modulation applied to a N0Sπ configuration. The complete set of stimulus types is listed in Table 1.

Figure 4 (left column) quantifies the effects of binaural modulation on the interaural disparity of a 100-Hz-wide band of noise centered around 500 Hz (a critical band). Three metrics of interaural disparity are shown as a function of modulation depth: normalized correlation (panel A), RMS of ITDs (panel B), and RMS of ILDs (panel C). The graphs are based on a numerical analysis of the stimuli computed as described in the "Appendix", Eqs. 11a

TABLE 1

Overview of stimulus conditions and the role of binaural cues

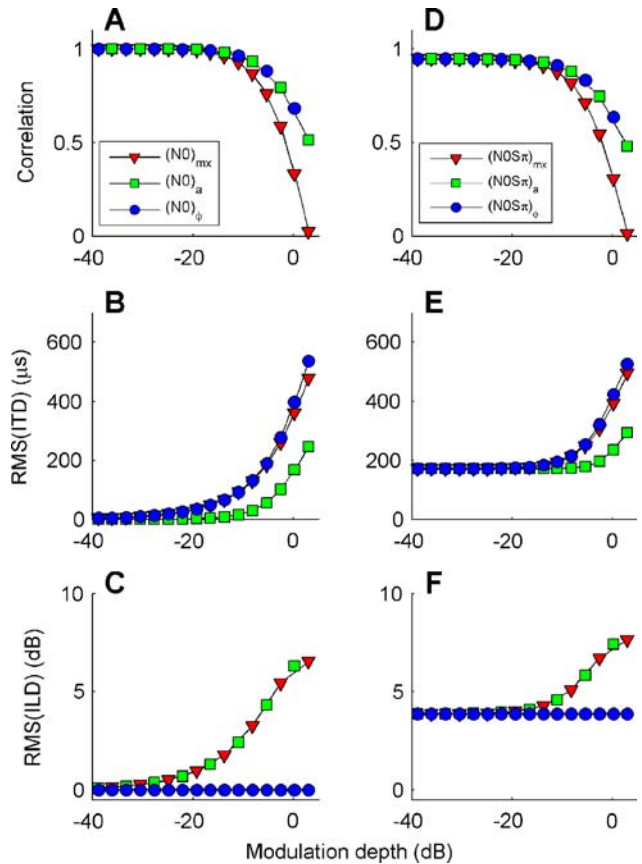| Stimulus	Modulation type	Binaural cues affected
$(N0S0)_d$	Diotic	None
$(N0S\pi)_d$	Diotic	None
$(N0S\pi)_{mx}$	Binaural mixed	Dynamic ITDs and ILDs
$(N0S\pi)_\phi$	Binaural QFM	Dynamic ITDs
$(N0S\pi)_a$	Binaural AM	Dynamic ILDs (spurious ITDs)

FIG. 4. The effect of depth of binaural modulation on three metrics of binaural disparity: normalized interaural correlation (**A**, **D**); RMS of dynamic ITD (**B**, **E**) and ILD (**C**, **F**). *Symbols* indicate the type of binaural modulation imposed on the stimulus: mixed (mx *triangles*), binaural AM (a *squares*), and binaural QFM (ϕ *circles*). In the left column (**A**, **B**, and **C**), the binaural modulation was applied to a diotic noise stimulus. In the right column (**D**, **E**, and **F**), the binaural modulation was applied to the sum of a diotic noise and an Sπ tone at –15 dB S/N ratio. In all cases, the binaural metrics were evaluated on a 100-Hz-wide noise band centered at 500 Hz.

and 11b. Instantaneous ILD and ITD were evaluated using the absolute value and angle of the complex-analytic stimuli; ITD was computed by multiplying the interaural phase by the period of a 500-Hz tone.

For the purposes of this study (see "Introduction"), the contrast between the binaural AM and binaural QFM is the crucial one because these two types of binaural modulation have the same effect on interaural correlation (squares and circles in Fig. 4A), but opposite effects on the fluctuations of ITDs and ILDs (Fig. 4B, C): binaural QFM affects only ITDs, while binaural AM primarily affects ILDs and has a much weaker effect on ITDs. The binaural mixed modulation combines the effects of the other two modulation types: For a given modulation depth, its decorrelating effect is larger than that of either binaural AM and QFM; its effect on ILDs equals that of binaural AM, whereas its effect on ITDs practically equals that of binaural QFM.

The differential effects of the binaural modulation types illustrated in the left column of Figure 4 are preserved when the modulation is applied not just to the N0 noise but to the whole N0Sπ complex. This is illustrated in the right column of Figure 4, showing the effects of binaural modulation on a diotic 100-Hz-wide band of noise to which an Sπ tone was added at a S/N ratio of −15 dB. The effects of binaural modulation on the noise + tone stimulus closely parallel the effects in the noise-alone case (Fig. 4, left column), except for the baseline binaural disparity at low modulation depths, which is caused by the Sπ tone.

Figure 1B shows how RMS(ITD) and RMS(ILD) covary when an N0 stimulus is subjected to the different types of binaural modulation. The binaural mixed modulation (diamonds) is seen to be "ILD/ITD neutral" in the sense that it produces virtually the same relative growth of ILDs and ITDs as does the addition of an Sπ signal (dashed line, replicated from panel A). Consistent with Figure 4, binaural QFM (stars) causes a growth of ITDs that is unaccompanied by a growth of ILDs, whereas binaural AM (crosses) causes an increase of ILDs accompanied with a relatively small increase of ITDs. The effect of binaural modulation on a mix of the diotic noise N0 and the antiphasic signal Sπ is illustrated by the arrows in Figure 1B. The magnitudes of ILDs and ITDs for an unmodulated stimulus (S/N=−13 dB) is indicated by the large circle on the dashed line. At small modulation depths, fluctuations of ILD and ITD are dominated by the presence of the Sπ signal, and binaural modulation has virtually no effect on their magnitudes (compare Fig. 4E, F). At larger modulation depths, binaural modulation starts to affect ITDs and/or ILDs. For binaural QFM, the effect is to increase RMS(ITD) without affecting RMS(ILD); this is indicated by the vertical arrow labeled "φ". For binaural AM, the effect is to primarily increase RMS(ILD), with only a slight effect on RMS(ITD); this is indicated by the horizontal arrow labeled "a". For the mixed binaural modulation, the effect is a combined growth of RMS(ILD) and RMS(ITD) along the dashed line; this is indicated by the arrowhead labeled "mx". Figure 1B quantifies the highly "cue-specific" scrambling effects of the binaural AM and the binaural QFM and the "cue-neutral" effect of mixed binaural modulation. The differential effects of binaural modulation on specific types of interaural disparity (Figs. 1, 4) provide the dissecting power needed to unravel the cues underlying binaural signal detection.

Stimuli

Detection of a 500-Hz signal was examined in the presence of a Gaussian noise ranging from 100 to 3,000 Hz, presented at a total level of 75 dB SPL, corresponding to a spectrum level of 40 dB. The 280-ms signals were temporally centered within the 300-ms maskers; both durations include 10-ms long $\cos^2$ on/off ramps. Independent noise tokens were randomly selected from a 5-s buffer for each presentation. A new noise buffer was computed for each experimental run.

Mixed modulation described in the previous section was imposed on both noise-alone ("reference") and noise-plus-tone ("target") intervals. The modulation frequency was 20 Hz, a value high enough to prevent the binaural system from "following" the fluctuations in binaural parameters (Grantham and Wightman 1978) and low enough for the type of modulation to be uncorrupted by peripheral filtering ("FM to AM conversion"; Blauert 1981). Modulation depth is expressed in decibels and equals 20log m, where m is twice the amplitude ratio of one sideband re the carrier. This definition generalizes the regular definition of AM depth. Computational details of the stimulus generation can be found in the "Appendix".

For each configuration of Table 1, modulation depths of −∞, −12, −7, −2, and 3 dB were tested. A modulation of −∞ dB means no modulation at all ($m=0$), resulting in conventional N0Sπ or N0S0 conditions. Note that even the 3-dB condition does not give rise to "overmodulation" in the sense of the sudden phase inversions that occur with AM when $m>1$. This is so because no value of m ever causes the envelope to be zero, as is clear from Figures 2 and 3. After visiting all 25 conditions once in random order, all conditions were visited again using a new randomization, and so on. Each condition was visited four times; reported thresholds are the average of the four estimates.

Apparatus, listeners, and procedure

Stimuli were generated digitally at a sample rate of 50 kHz, played via a D/A converter (Tucker-Davis Technologies PD1), low-pass-filtered at 20 kHz, and presented via TDH 39 headphones to listeners seated in a double-walled soundproof cabin. Five female students with normal hearing, 21 to 25 years old, served as listeners. Each listener received several hours of training prior to participating in the experiments.

The stimuli were presented in a two-interval, two-alternative forced choice task. Each trial consisted of a 500-ms warning interval followed by two 300-ms observation intervals separated by 400 ms. Intervals were marked by a computer monitor. The signal was presented with equal a priori probability in either the first or the second interval. The listener had to indicate the signal interval; correct-answer feedback was provided visually after each response. The level of the signal was varied adaptively according to a two-down one-up rule in order to estimate 70.7% correct. After two reversals, the initial 3-dB step size was reduced to its final size, 1.5 dB. A run was terminated

after 10 more reversals. Estimates of the thresholds result from averaging the signal levels at the last eight reversals.

RESULTS

Figure 5 shows the thresholds of the individual listeners (panels A–E). Error bars indicate ±one standard deviation of the four estimates. Stimulus configurations (see Table 1) are indicated in the graph. The data obtained with the $-\infty$ dB modulation depth serve as a reference: They are conventional N0Sπ and N0S0 thresholds. The absence of modulation renders the four binaural conditions identical; any differences between the corresponding four thresholds reflect measurement variability. The MLDs ranged from 8 to 13 dB across listeners.

The $(N0S0)_d$ and $(N0S\pi)_d$ thresholds (solid black symbols) do not show a systematic variation with modulation depth. (This was tested with a two-way analysis of variance using modulation depth and listener as factors, yielding $p > 0.1$ and $p > 0.25$ for the main effect of modulation depth in the $(N0S0)_d$ and $(N0S\pi)_d$ conditions, respectively.) Because these are the conditions in which the modulation is diotic, this implies that modulation per se, although clearly audible, does not affect the monaural and binaural thresholds of the 500-Hz tone. This strongly suggests that the *cues* underlying monaural and binaural detection of the tone are unaffected by the modulation as long as the modulation does not introduce any binaural disparities. (Recall that modulation was simultaneously applied to the noise and the tone, leaving the interaction between noise and tone unaffected.) That diotic modulation does not have a clear effect on detection justifies an interpretation of any effects in the remaining configurations solely in terms of the *binaural* aspects of the modulation.

In contrast to the diotically modulated conditions, the $(N0S\pi)_{mx}$ thresholds (downward triangles) show a clear effect of modulation depth m. With increasing m, thresholds grow from a minimal, N0Sπ value and approach the N0S0 threshold at the highest value of m, 3 dB. Thus the 20-Hz mixed binaural modulation, when strong enough, causes a substantial reduction of the binaural advantage. The modulation in this case is antiphasic in both its QFM and AM constituents, resulting in a decorrelation of the waveform that is mediated by both dynamic ITDs and ILDs (Figs. 1, 4). In this respect, $(N0S\pi)_{mx}$ resembles a classic NρSπ condition in which the interaural correlation ρ of the noise is varied by mixing independent noise sources (Robinson and Jeffress 1963). The connection between $(N0S\pi)_{mx}$ and NρSπ thresholds will be elaborated in the next section.
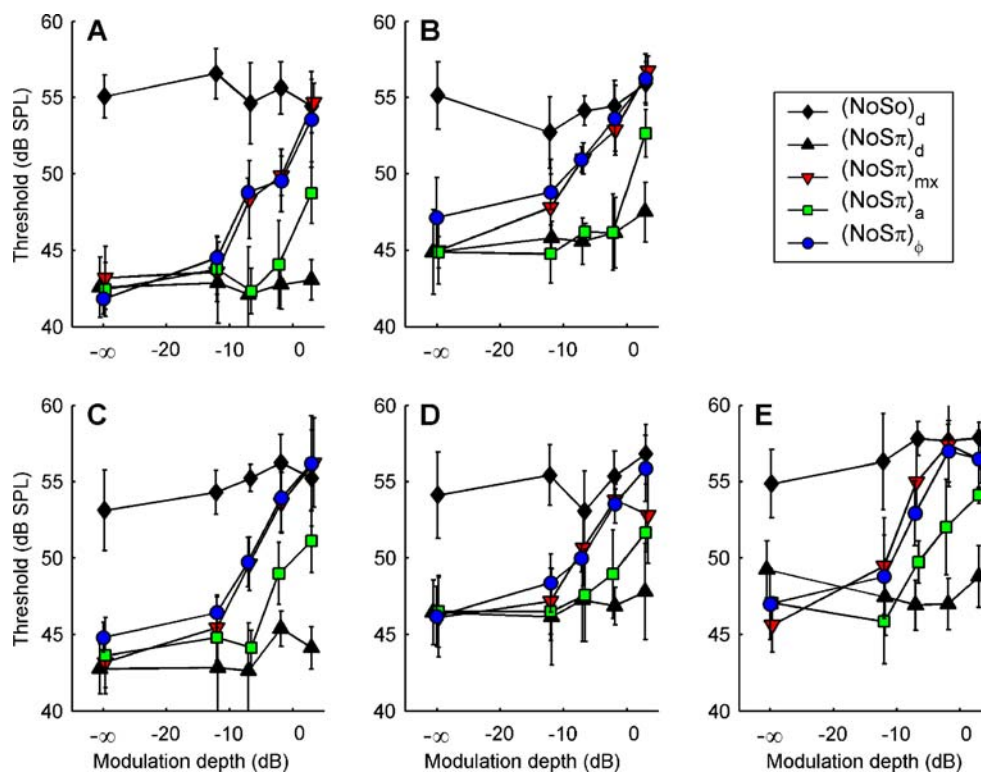
FIG. 5. Thresholds for detecting a 500-Hz tone in the presence of wideband noise. *Each symbol* is the average of four thresholds; *error bars* indicate ±one standard deviation. Each panel displays the data from one listener. *Symbols* indicate different conditions. *Solid filled symbols* are used for the control conditions, in which diotic modulation was applied to N0S0 (*solid black diamonds*) and N0Sπ (*solid black triangles*). *Remaining symbols* indicate different types of binaural modulation applied to N0Sπ: mixed (*downward triangles*), binaural AM (*squares*), and binaural QFM (*circles*).

Next consider the conditions in which binaural modulation is imposed in either an ILD- or an ITD-specific way: the $(N0S\pi)_a$ and the $(N0S\pi)_\phi$ conditions, respectively. The $(N0S\pi)_a$ thresholds (open squares) show a weak effect of modulation depth: only at the two highest values of modulation depth (−2 and 3 dB) are the thresholds elevated by the binaural AM. In contrast, the $(N0S\pi)_\phi$ curves (open circles) practically coincide with the $(N0S\pi)_{mx}$ curves. (This was confirmed by a three-way analysis of variance with factors modulation type, listener, and modulation depth, yielding $p > 0.25$ for the main effect of modulation type.) Thus, binaural QFM has a much more pronounced effect on binaural detection than binaural AM, implying that dynamic ITDs are more important than dynamic ILDs in the N0Sπ detection task. The overlap of $(N0S\pi)_{mx}$ thresholds (in which ITDs and ILDs are equally manipulated) and $(N0S\pi)_\phi$ thresholds (in which only ITDs are manipulated) strongly suggest that dynamic ITDs completely determine the low-frequency N0Sπ condition used in the present study and that ILDs do not play any role.

To illustrate the dominance of ITDs more quantitatively, consider the data of listener A, whose unmodulated N0Sπ threshold is ~42 dB SPL. In terms of the metrics of interaural disparity (Fig. 1), the Sπ signal at threshold produces RMS(ITD) = 130 μs and RMS(ILD) = 3.4 dB. Binaural modulation interferes with these metrics in the following way. In the $(N0S\pi)_{mx}$ condition at −7 dB modulation depth, the modulation introduces interaural disparities in the noise-only stimulus amounting to RMS(ITD) = 150 μs and RMS(ILD) = 3.8 dB (Fig. 4B, C). In both the ITD and ILD domains, the modulation can therefore be expected to seriously interfere with the subtle effects of adding the 42-dB Sπ tone to unmodulated noise. Indeed, the $(N0S\pi)_{mx}$ threshold at −7 dB modulation depth shows a sizeable loss of MLD: The threshold is elevated by 6 dB. We next examine whether this threshold elevation is due to a scrambling of the ITD cue, the ILD cue, or both. To answer this question, consider the $(N0S\pi)_\phi$ and $(N0S\pi)_a$ thresholds obtained at the same modulation depth, −7 dB. The $(NoS\pi)_\phi$ condition introduces exactly the same baseline RMS(ITD) as the $(N0S\pi)_{mx}$ condition (150 μs), but does not evoke any baseline ILDs (Fig. 4C). Conversely, the $(NoS\pi)_a$ condition introduces only a small ("spurious") baseline RMS (ITD) of ~40 μs and the same baseline RMS(ILD) as the $(NoS\pi)_{mx}$, i.e., 3.8 dB. Now observing that both the $(NoS\pi)_\phi$ and $(NoS\pi)_{mx}$ thresholds are elevated by 6 dB, while the $(NoS\pi)_a$ is not elevated at all, the dominance of ITDs is evident. (A more detailed analysis, which also considers the effect of modulation on signal + noise stimuli, cf. Figure 1B, is presented in the next section.)

The occurrence of spurious ITD modulation in the $(N0S\pi)_a$ condition (squares in Fig. 4B) leaves open the possibility that the effect of modulation in this condition is in fact caused by spurious ITD modulation, not the "proper" ILD modulation. If that is true, the binaural amplitude modulation itself has no effect at all, not even at the highest values of m, and the dominance of dynamic ITDs in the N0Sπ detection task would be complete. The evaluation of that possibility requires a quantitative analysis of interaural correlation and dynamic ITDs and ILDs.

Control experiment: varying modulation frequency

Our choice of a 20-Hz modulation rate was based on the assumption that the auditory filter at 500 Hz is too wide to interfere with the phases of the stimuli, i.e., that there will be no FM-to-AM conversion (or vice versa) that can spoil the ITD- or ILD-specific character of the binaural modulation. The observed contrast between the thresholds obtained with different binaural modulation types already shows that auditory filtering does preserve the character of the modulation to a high degree. To further analyze the effect of peripheral filtering on binaural modulation, listeners A and B were tested in a number of additional $(N0S\pi)_a$ and $(N0S\pi)_\phi$ conditions using modulation rates of 10, 20, 40, 80, and 160 Hz. The modulation depth was −2 dB, a value for which a sizeable contrast between $(N0S\pi)_a$ and $(N0S\pi)_\phi$ thresholds was observed in the main experiment (Fig. 5). Procedures were identical to those of the main experiment; all conditions were randomized.

Our expectations were as follows. Modulation rates greatly exceeding the critical bandwidth will cause a loss of modulation character: FM-to-AM conversion is expected to effectively turn any binaural modulation into a mixed binaural condition. In particular, supracritical modulation rates should cause a convergence between $(N0S\pi)_a$ and $(N0S\pi)_\phi$ thresholds.

The data (Fig. 6) confirmed our expectations. The contrast between $(N0S\pi)_a$ and $(N0S\pi)_\phi$ conditions previously observed at the 20-Hz modulation rate is also present at 10 Hz. At the 40-Hz modulation rate, the contrast is reduced for listener B, but not for listener A. For both listeners, the contrast between $(N0S\pi)_a$ and $(N0S\pi)_\phi$ conditions is reduced at 80 Hz, and disappears at 160 Hz. Bearing in mind that at a modulation rate f_{mod} the sidebands of a given stimulus component are $2f_{mod}$ apart, the data are consistent with a critical bandwidth of about 100 Hz, in agreement with previous estimates of critical bandwidth from binaural detection experiments (Van der Heijden and Trahiotis 1998). At the low end of the modulation rates tested, note that at 10 Hz the $(N0S\pi)_\phi$ thresholds are still elevated re the unmodulated N0Sπ threshold. Apparently, 10 Hz is still too fast for the binaural system to track the dynamic binaural disparities evoked by the binaural modulation. This is consistent with data from Grantham and Wightman (1978).
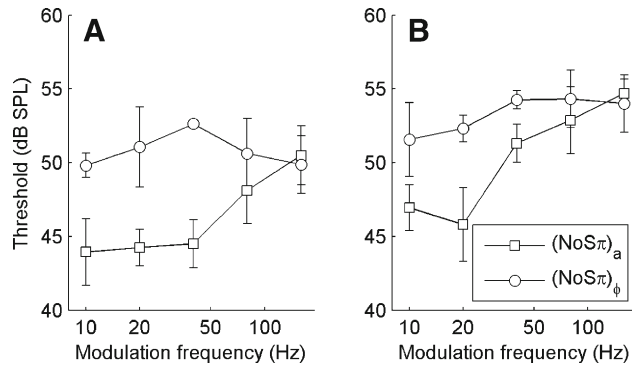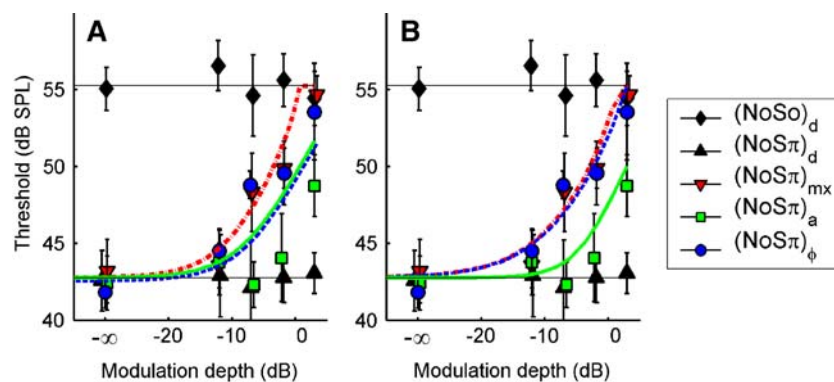
FIG. 6. A, B Thresholds from an additional experiment in which modulation rate was varied at a fixed modulation depth of −2 dB. Only two types of binaural modulation were used: (N0Sπ)$_a$ (*open squares*) and (*open circles*). Error bars indicate ±one standard deviation. Each panel displays the data from one listener.

QUANTITATIVE ANALYSIS OF THE DATA

The patterning of the data from the main experiment (Fig. 5) is systematic, and the main observations have already been stated. The more quantitative analysis pursued in the present section will not alter the main conclusions. Yet it will prove interesting to explore how well the quantitative details of the data can be captured by simple models, which may serve as a starting point for full-fledged models that explain binaural signal detection in terms of processing of ITDs and ILDs.

Failure of stimulus interaural correlation

Under the assumption that interaural correlation is the sole determinant of binaural detection, the effects of binaural modulation are mediated by its effect on the correlation of the noise and of the noise-plus-tone complex. In particular, for a given value of *m*, the (N0Sπ)$_a$ and (N0Sπ)$_\phi$ thresholds should be identical

because their stimuli are equivalent in terms of interaural correlation (Fig. 4A, D). Although the failure of this approach is evident from the data (Figs. 5, 6), it is instructive to examine the quantitative details.

The effect of binaural modulation on the correlation of the masker and the masker + signal combination was analyzed in quantitative detail (see "Appendix"). This allowed a comparison between our thresholds with the corresponding NρSπ thresholds, in which the correlation ρ of the masking noise is varied by mixing independent noise sources (Robinson and Jeffress 1963). NρSπ data are not available for the listeners participating in this study, but they can be predicted very well from the N0Sπ and N0S0 thresholds by parsimonious theoretic arguments (Durlach 1972; Van der Heijden and Trahiotis 1997). Predictions of the thresholds were obtained by bandpass filtering the modulated noise maskers (100-Hz-wide critical band centered at 500 Hz) and computing the crosscorrelation of the filtered noise waveforms. These values of crosscorrelation were used to predict the thresholds; computational details can be found in the "Appendix".

The correlation-based predictions are shown in Figure 7A for listener A. The (N0Sπ)$_{mx}$ predictions (dash-dotted line) match the data reasonably well. This is not surprising because the (N0Sπ)$_{mx}$ condition contains an "equal mix" of ITDs and ILDs, rendering it most similar to NρSπ. The (N0Sπ)$_a$ predictions (solid line) show a smaller effect of binaural modulation than the corresponding (N0Sπ)$_{mx}$ predictions. That trend is also present in the data, but the predicted contrast between (N0Sπ)$_a$ and (N0Sπ)$_{mx}$ is too small. As anticipated, the predictions completely fail to predict the difference between (N0Sπ)$_\phi$ and (N0Sπ)$_a$ thresholds. The two sets of predictions are identical (solid line obscures gray dashed line) because the effect of binaural modulation on correlation is the same for the

FIG. 7. Predictions based on interaural correlation in a 100-Hz-wide noise band centered at 500 Hz (see text). *Symbols* are the data from listener A, replotted from Figure 5. *Line styles* and *symbols* indicate the type of binaural modulation: mixed (*dash-dotted line, downward pointing triangles*); binaural AM (*solid line, squares*), and binaural QFM (*dashed line, circles*). *Black symbols and horizontal lines* are the control conditions as in Figure 5. In **A**, interaural correlation was computed directly from the filtered waveforms. In **B**, the filtered waveforms were subjected to a compressive power law (0.25 dB/dB) prior to computing the correlation. Note that the *solid and dashed lines* in **A** coincide, because the predictions for (N0Sπ)$_a$ and (N0Sπ)$_\phi$ are identical. In **B**, on the other hand, the *dashed* and *dash-dotted lines* largely overlap, because the predictions for (N0Sπ)$_\phi$ and (N0Sπ)$_{mx}$ are very similar.

two configurations (Fig. 4A, D). In contrast, the data reveal a unmistakable contrast between $(N0S\pi)_{\phi}$ and $(N0S\pi)_a$, and show instead an equivalence between $(N0S\pi)_{\phi}$ and $(N0S\pi)_{mx}$ thresholds. The model also fails to reproduce the latter equivalence: It predicts higher thresholds in the $(N0S\pi)_{mx}$ configuration.

In sum, the predictions based on interaural correlation are in the proper range, but miss systematic trends in the data.

Interaural correlation after waveform compression

The observed asymmetry between ITDs and ILDs might be reproduced by compressing the effective waveforms from the two ears prior to computing the correlation. The idea is that nonlinear waveform compression will reduce the envelope fluctuations in each ear, causing a reduction of the dynamic ILDs. This scheme was first proposed by Van de Par and Kohlrausch (1998) and further pursued by Bernstein et al. (1999). Because ITDs are not affected by the compression, the interaural correlation computed from compressed waveforms is more sensitive to ITDs than to ILDs of the original (uncompressed) waveforms.

The threshold predictions of the previous section were repeated, this time applying an instantaneous nonlinear compression of the filtered waveforms according to the power law $y(t) = x(t)|x(t)|^{\alpha-1}$. The compression parameter α determines the steepness of the I/O curve: The output/input ratio is α dB/dB, with $0<\alpha<1$. Figure 7B shows the predictions for $\alpha=0.25$ dB/dB together with the data of listener A. The predictions are much better than those of Figure 7A. As anticipated, waveform compression increases the sensitivity to ITD jitter compared to ILD jitter, leading to a difference between $(N0S\pi)_{\phi}$ and $(N0S\pi)_a$ threshold predictions that matches the data well. The predictions of the $(N0S\pi)_{mx}$ and $(N0S\pi)_{\phi}$ thresholds are almost identical, again consistent with the data. Note that the choice $\alpha=0.25$ dB/dB corresponds to a strongly compressive nonlinearity. Less compressive nonlinearities (larger α) fail to reproduce the observed equivalence between $(N0S\pi)_{mx}$ and $(N0S\pi)_{\phi}$ thresholds. In contrast, more compressive nonlinearities ($\alpha<0.25$) produce a match with the data that is comparable to the choice $\alpha=0.25$. Because increasingly smaller values of α will reduce ILDs more and more, the success of smaller choices of α suggests that the dominance of ITDs over ILDs is complete.

Webster revisited

Webster (1951) hypothesized that detection of an Sπ tone in a wideband N0 noise requires the time-varying ITD to exceed a certain critical value during the stimulus presentation. The present data do confirm the

dominance of dynamic ITDs proposed by Webster, but at this stage, it is not clear whether the patterning of Figure 5 can be explained solely on the basis of dynamic ITDs, or whether a properly weighted mix of ITDs and ILDs is needed to account for the effects of binaural modulation.

In order to address this question, a slight variation on Webster's hypothesis is introduced: The criterion for binaural signal detection is a *sufficient change in the RMS of the ITD*. In a full-fledged model, a "sufficient change" is judged on the basis of statistical significance, the assessment of which requires a statistical analysis of the time-varying ITD of the bandpass-filtered stimulus over a finite stimulus duration. The present numerical analysis is less ambitious and is limited to the expected values of the RMS of the ITD.

The time-varying ITD of our stimuli was computed as follows. After restricting the complex-analytic stimuli to a 100-Hz-wide band around 500 Hz, the instantaneous interaural phase difference was converted to ITD by multiplication by 2 ms, the period of 500 Hz. For each of the stimulus configurations $(N0S\pi)_{\phi}$, $(N0S\pi)_a$, and $(N0S\pi)_{mx}$, and for each modulation depth, the RMS of the ITD was computed for 1000 independent 1-s waveforms. Recall that many of our noise-alone conditions already contain sizeable ITD fluctuations caused by the binaural modulation of the stimuli. Thus, the mere size of ITD fluctuations in the noise + tone condition is not a useful metric for detectability; instead, one has to compare the magnitude of ITD fluctuations between the noise-alone (reference) stimulus and the noise-plus-signal (target) stimulus at threshold.

Figure 8 shows this comparison. Each symbol relates RMS(ITD) for a given reference stimulus (abscissa) to RMS(ITD) for the corresponding target stimulus at threshold (ordinate). Different symbols represent the different modulation types as indicated in the graph. Each panel represents data of a single listener; panel arrangement and use of symbols are the same as in Figure 5. Error bars were derived from the standard deviation (SD) of the threshold estimates (the error bars in Fig. 5) by computing the RMS(ITD) at threshold ±SD.

One may view Figure 8 as an alternative way of representing the thresholds of Figure 5. In Figure 5, the amount of jitter was expressed in terms of modulation depth (abscissa of Fig. 5); in Figure 8, the amount of jitter is expressed in terms of the dynamic ITD of the reference condition, i.e., RMS(ITD). In Figure 5, binaural detection was expressed in terms of signal level at threshold (ordinate); in Figure 8, the binaural detection is expressed in terms of the value RMS(ITD) produced by adding to the threshold-level signal to the noise.

The most important characteristic of Figure 8 is the convergence of different modulation types compared
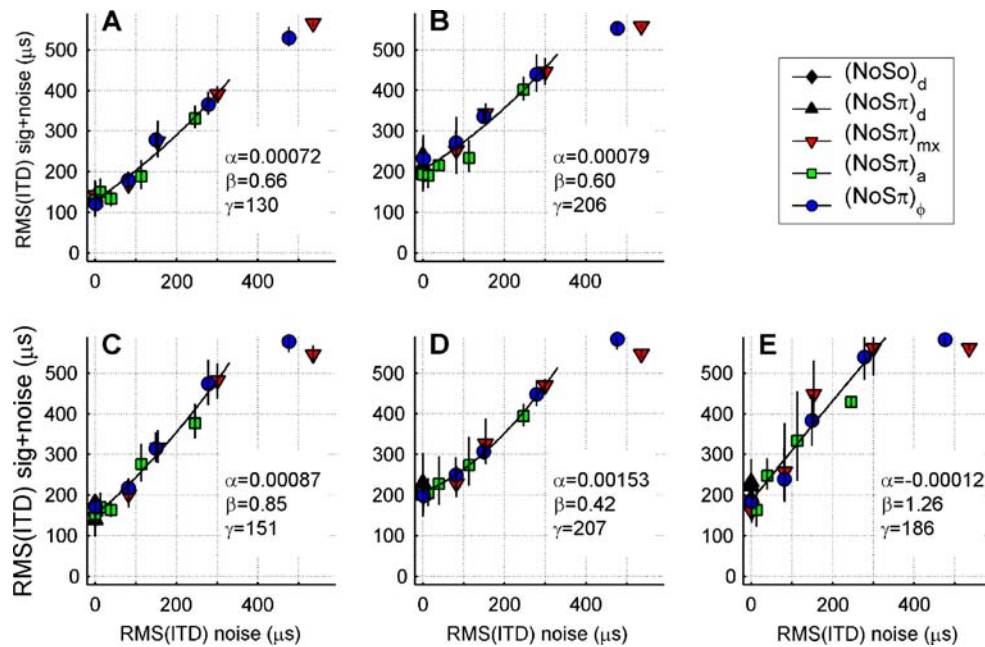
FIG. 8. Representation of the detection thresholds of Figure 5 in terms of the size of ITD fluctuations, RMS(ITD), in a 100-Hz-wide band centered at 500 Hz. For each detection threshold the relation is shown between the RMS(ITD) of the noise-alone condition (*abscissa*) to the same metric extracted from the noise + signal condition at the detection threshold (*ordinate*). *Symbols* as in Figures 5 and 7. *Lines* are second order polynomial fits $y = \alpha x^2 + \beta x + \gamma$ to the data. Fit parameters are indicated in each plot. The $(N0S\pi)_{mx}$ and $(N0S\pi)_{\phi}$ thresholds obtained with 3-dB modulation depth were excluded from the fits, because their MLD vanished.

to Figure 5: The different curves in Figure 5 collapse onto a single curve in Figure 8. Within the measurement error, detectability of the signal is well described by a single-valued relation between RMS(ITD) in the reference and target conditions. (The precise shape of that relation is of secondary importance.) Different modulation types yield different detection thresholds, but these differences largely disappear when the thresholds are converted from signal level (Fig. 5) to magnitude of ITD fluctuations (Fig. 8). This identifies RMS(ITD) as a reliable predictor of binaural detection in the conditions reported here. Note that the $(N0S\pi)_a$ thresholds (open squares) fit well in this ITD-based description. This is consistent with the notion that the effect of binaural AM on binaural detection is mediated by spurious ITDs and not by ILDs.

The numerical relation between RMS(ITD) with and without the signal was quantified by second-order polynomial regression (Fig. 8, lines). The coefficients of the fits for each listener are indicated in the graph. The datapoints with negligible MLD (the two points in each panel for which RMS(ITD) > 400 μs in the noise-alone condition) were excluded from the fits because the detection in these cases does not require binaural processing (the exluded points are the $(N0S\pi)_{\phi}$, and $(N0S\pi)_{mx}$ thresholds obtained at maximal modulation depth, 3 dB; see Figure 5).

A simple model is based on the crude approximation of the curves in Fig. 8 by straight lines having unity slope and intercept Δ. This simplification

corresponds to the following heuristic criterion for binaural detection: *the addition of the signal to the masker is just detectable when it increases the RMS of the ITD by a critical amount Δ.* This criterion differs in two respects from Webster's (1951) hypothesis ("ITD has to exceed a threshold value"). First, it uses RMS rather than threshold value as a metric of ITD-based interaural disparity. Second, it generalizes Webster's formulation from conditions with a diotic reference stimulus (no ITDs) to situations in which the reference stimulus already contains nonvanishing dynamic ITDs. Yet despite these technical differences, the two criteria express precisely the same proposition, namely, that binaural detection can be understood from the dynamic processing of ITDs.

Figure 9 shows the predictions of the thresholds based on the above hypothesis, using Δ > 130 μs. For each stimulus condition, the predicted thresholds are those signal levels which yield an increase in the RMS of the ITD by the critical amount Δ. "Negative MLDs" were excluded: Any threshold predictions that exceeded the N0S0 threshold were replaced by the N0S0 threshold. For reference, the data of listener A are shown as symbols (cf. Fig. 5A). The predictions match these data very well. Obviously, the data of some of the other listeners, whose ITD statistics deviate more from a straight line of unity slope (e.g., Fig. 8D), will be predicted less accurately by the simple model. The significance of the model, however, is not in its success in predicting a single
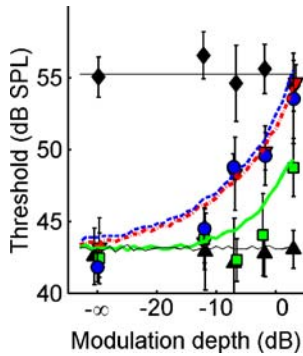
FIG. 9. Predictions derived from a Webster-type model, requiring a 130-μs increment of RMS(ITD) in a 100-Hz-wide band around 500 Hz for binaural detection (see text). *Symbols* are the data from listener A, replotted from Figure 5. *Line styles and symbols* indicate binaural modulation type as in Figure 7.

listener's data, but in its ability to reproduce systematic trends in the data. In that respect, the model is successful: It correctly predicts the equivalence of $(N0S\pi)_{mx}$ and $(N0S\pi)_\phi$ thresholds, the difference between $(N0S\pi)_\phi$ and $(N0S\pi)_a$ thresholds, and the growth of all three types of thresholds with modulation depth. The success of the model is remarkable given its extreme simplicity; it attempts to capture binaural detection in a single parameter Δ.

In sum, the analysis of ITD statistics (Fig. 8) suggests that the binaural detection thresholds can be explained exclusively in terms of dynamic ITDs and that ILDs do not play a role in a N0Sπ condition. The growth of $(N0S\pi)_a$ thresholds with modulation depth can be explained by the spurious dynamic ITDs occurring with the binaural amplitude modulation used here. A heuristic model (Fig. 9) inspired on Webster's (1951) hypothesis reproduces the essence of the data for all modulation types and depths.

DISCUSSION

In this study, we applied a binaural modulation technique that introduces binaural disparities consisting of either dynamic ITDs, ILDs, or both. The technique was used to impose specific types of jitter on the stimuli of a binaural detection task. It was found that binaural detection of a 500-Hz tone in wideband noise is more sensitive to the introduction of ITD-specific jitter of the stimulus than it is to ILD-specific jitter. The asymmetry between ITDs and ILDs clearly indicate that interaural correlation of the stimulus waveform is not a good predictor of binaural detection in this N0Sπ task. Quantitative analysis showed that the data are in agreement with the hypothesis that detection is solely based on the dynamic ITDs in the stimulus and that ILDs do not play a role. This finding can either be understood in

terms of a dynamic tracking of fluctuating ITDs or in terms of interaural correlation computed from strongly compressed versions of the monaural inputs. The former view corresponds to a lateralization model of MLD, in which the fluctuating ITD is accessible to the binaural system and subject to rapid dynamic tracking or at least to some form of statistical analysis. The latter view explains the dominance of ITDs in terms of a neural representation of the monaural inputs that is "compressed" in the sense that it is relatively insensitive to intensity fluctuations. The combination of such a compressive stage and a straightforward binaural coincidence detector can explain our findings without the need for analyzing dynamically varying binaural parameters.

It should be noted that these two alternatives are not necessarily mutually exclusive and that intermediate modes of processing are conceivable. For instance, in a Webster-type lateralization model, the magnitude of ITD fluctuations may well be evaluated in terms of a rectify + integrate scheme, thus dispensing with the need for dynamic tracking of ITD. Conversely, in compression + correlation schemes, the correlation may well be dynamically tracked with a certain temporal resolution. The "dynamic versus integrated" contrast between the two models should therefore not be exaggerated. It seems to us that a more fundamental contrast between the two approaches is the stage of processing at which the ITD dominance is realized: In lateralization models, it happens after binaural interaction, while in compression + correlation schemes, it happens in the monaural pathways, i.e., prior to binaural convergence. Again one can conceive of a mixture of both processing schemes, but at least the monaural/binaural contrast is amenable to physiological testing, as will be discussed below.

Comparison with the multiplied-noise study of Van de Par and Kohlrausch

Our finding that ITDs dominate binaural detection appears to contradict those of Van de Par and Kohlrausch (1998; abbreviated as vdP&K below). In their low-frequency, narrowband data (Fig. 4, restricted to CF ≤ 1 kHz), they observed no systematic effect of the angle between the signal and the multiplied-noise masker. The authors concluded from their low-frequency data: "It seems that at low frequencies, the sensitivity to dynamically varying IIDs or ITDs can be singly related to changes in the interaural correlation of the waveforms. Thus there seems to be no need to separately make assumptions about the sensitivity to changes in IIDs and ITDs." This appears to contradict the asymmetric roles of ITDs and ILDs reported in the current study. There

are, however, major differences in methodology and interpretation between their study and ours.

In order to be able to control the relative phase of signal and masker, vdP&K used "multiplied noise", i.e., a tone multiplied by a low-pass Gaussian noise. The statistics of the fine structure and the envelope of multiplied noise are very different from those of Gaussian noise, and this is known to affect their masking potency (Van der Heijden and Kohlrausch 1995). Another important factor is masker bandwidth. The very narrow (25-Hz-wide) noise maskers used by vdP&K cause slowly varying interaural disparities that are more likely to allow dynamic binaural processing than the wideband noise used in the present study. More generally, there are large qualitative and quantitative differences between narrowband and wideband noise maskers. This is true for monaural detection (Richards 1992) and for binaural detection: the size of MLD, for instance, which is typically 12 dB in wideband conditions used in the present study, exceeds 25 dB in several of the narrowband conditions employed in vdP&K.

Importantly, the multiplied-noise paradigm used by vdP&K gives rise to an artifact that renders the interpretation of their low-frequency data somewhat problematic. The authors reason that adding the Sπ signal in phase to the masker ("$\vartheta > 0$" condition) produces exclusively interaural amplitude differences; interaural phase is assumed to stay unaffected. This, however, is not true in the neighborhood of the sharp minima in the noise envelope. Each zero crossing of the low-pass noise used to generate the masker results in a segment of time during which the amplitude of the Sπ signal exceeds that of the noise. As illustrated in Figure 10, the noise + tone stimulus is *antiphasic* during those segments. Thus the stimulus that was designed to yield only ILD cues in fact contains interaural phase reversals. The authors incidentally mentioned this imperfection in the description of their methods, but did not analyze its potential consequences. More recent work (Bernstein et al. 2001; Boehnke et al. 2002), however, has shown that human listeners are very
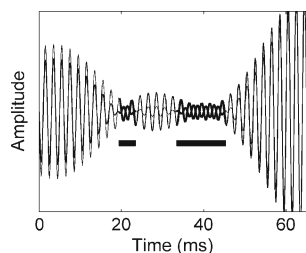
FIG. 10. Phase-reversal artifact in the multiplied-noise masking paradigm used by Van de Par and Kohlrausch (1998). An Sπ tone was added to a 25-Hz-wide multiplied-noise masker at an S/N ratio of −20 dB. Left- and right-ear waveforms are shown superimposed. The relative phase of masker and signal was chosen with the intention to produce only ILDs, but in fact gives rise to antiphasic segments (*thick lines*).

sensitive to brief segments of interaural disparity: Antiphasic segments as short as 1.4 ms can be detected when embedded in diotic noise. These findings prompt a closer look at the potential effects of the phase inversions in the stimulus conditions used by vdP&K.

When adding an Sπ tone at −20 dB S/N level to a 300-ms long, 25-Hz-wide multiplied-noise band, antiphasic segments of 2 ms or longer occur in 99.9% of the cases; an average of six such segments occur in each 300-ms presentation. One or more antiphasic segments lasting at least 5 ms occur in 79% of the cases. When reducing the S/N ratio to −25 dB (a value close to the most sensitive binaural thresholds in vdP&K), one or more phase-reversed segments lasting at least 2 ms still occur in 89% of the cases. In conclusion, the unintended phase cues are likely to have played a role in the thresholds reported by vdP&K, and this renders the multiplied-noise paradigm unsuitable for assessing the relative contributions of ILDs and ITDs to low-frequency binaural detection. Thus, despite appearances, the low-frequency data of vdP&K are compatible with dynamic ITDs dominating binaural detection.

Binaural processing of dynamic ITDs

The first approach to explain the dominance of ITDs in terms of processing strategies or physiological mechanisms is to emphasize the *dynamic* character of binaural processing. This approach originates with Webster (1951), who described ITD (or effective ITD within a critical band) as a quantity that fluctuates during the stimulus and furthermore assumed that these fluctuations are accessible to the binaural system. Webster proposed that detection of the stimulus merely requires the fluctuating ITD to exceed a certain threshold at an arbitrary instant of time. This approach was carried further by Jeffress et al. (1956) and Hafter and Carrier (1970), who also took ILDs into consideration. The overall aim of this classic work is a synthesis of binaural detection and lateralization ("lateralization model of MLD").

The use of wideband stimuli causes the fluctuations in interaural disparities to be quite rapid (with rates on the order of a critical bandwidth). For narrowband stimuli, the fluctuations are slower (with rates on the order of the stimulus bandwidth). Thus the interaural fluctuations get faster with increasing stimulus bandwidth. If binaural processing has limited temporal resolution, one would expect performance to decline with bandwidth. Such a decline of binaural performance with stimulus bandwidth is indeed generally observed (e.g., Zurek and Durlach 1987). For narrow bandwidths, it is entirely reasonable to postulate the dynamic processing of interaural disparities—if only because dichotic narrowband stimuli simply sound like they are "moving around". For wideband stimuli, on the other hand, the

assumption of dynamic processing is more problematical because reports of "binaural sluggishness" (Grantham and Wightman 1978) strongly suggest that dynamic processing of interaural disparities breaks down at rates as low as 10 Hz. It is difficult to reconcile such a degree of sluggishness with Webster's dynamic ITD processing in the case of wideband stimuli as in the present study.

Several behavioral and physiological studies, however, suggest a degree of sluggishness that is less severe than indicated by the findings of Grantham and Wightman (1978). The detectability of very brief segments of binaural disparity within an otherwise diotic stimulus (Bernstein et al. 2001; Boehnke et al. 2002) was already mentioned. These studies in human listeners were inspired by a behavioral study in the barn owl by Wagner (1991), showing an exquisite sensitivity to very brief (1-ms) jumps in ITD of a wideband noise. Wagner (1992) also presented physiological data from the inferior colliculus (IC) of the barn owl, reporting sensitivity of binaural "broad-band neurons" in the IC to such brief jumps in ITD. Binaural cells in the IC of the cat are able to "track" sinusoidal variations in interaural correlation at rates well over 100 Hz (Joris et al. 2006). Additional psychophysical and physiological data in support of fast binaural processing were reported by Siveke et al. (2008). Although the human ability to detect oscillating correlation breaks down at much lower rates (~10 Hz; Grantham and Wightman 1978), such findings warn against an overly rigid interpretation of binaural sluggishness (see also Stellmack et al. 2005; Thompson and Dau 2008).

Compression + correlation schemes

The dominance of dynamic ITDs observed in the present study does not in itself prove that the binaural system is able to track the dynamically varying ITD of the (effective) stimulus. As an alternative explanation, it was shown (Fig. 7B) that interaural correlation is a good predictor of the binaural thresholds in this study, provided that interaural correlation is not directly computed from the effective stimulus, but from a strongly compressed version of the waveforms. This mode of modeling the data depends on an *integrated* stimulus metric ("interaural correlation after monaural compression"), not requiring a continual tracking of rapidly fluctuating binaural disparity as in the Webster approach. Stimulus-oriented modeling of binaural processing involving a compressive stage followed by crosscorrelation has been considered previously by Van de Par and Kohlrausch (1998) and by Bernstein et al. (1999). A physiological realization of this scheme would be based on the spike count over the entire stimulus duration at the output of a binaural coincidence detector—provided that the monaural inputs to this detector are in some sense a compressed version of the stimulus waveforms.

What does it mean for a neural response to carry a "compressed representation" of a sound waveform? It is important to realize that the neural response to sound is not an analog waveform, but a time series of equal-amplitude action potentials, whose only relation to the stimulus waveform is their timing. In that sense, the neural response of single neurons is already a perfectly compressed version of the stimulus. On the other hand, envelope fluctuations are still reflected in the neural response by way of corresponding fluctuations in instantaneous firing probability (Joris and Yin 1992, 1998), and this envelope coding does not show a compressed character at the level of the monaural afferents. Despite the limited dynamic range of most AN-fiber responses to pure tones, envelope coding by low-frequency AN fibers is rather linear (Joris and Yin 1992; Fig. 2). Bushy cells in the cochlear nucleus (the inputs to the phase-sensitive binaural cells in the olivary complex) show an even smaller dynamic range in response to pure tones than AN fibers (Joris et al. 1994). This is suggestive of a compressive transformation between AN fibers and bushy cells, but it is presently unclear whether this compressed dynamic range in response to static tones also carries over to dynamic intensity fluctuations, i.e., envelope coding. Reduced dynamic range of single-tone responses can at most provide indirect support for the type of compressive model considered here.

The highly nonlinear, discrete, and stochastic nature of neural responses makes it difficult to attach a precise meaning to the question of whether "the monaural responses are compressed". The analog sound waveform and its neural response are simply too dissimilar for one to be a compressed version of the other. In the context of the present study, a more precise and operational question is: When the monaural responses are fed to a cross-coincidence detector, is the output of the detector more sensitive to decorrelations carried by ITDs than to equally large decorrelations carried by ILDs? If so, the circuit effectively behaves like an analog compress + correlate model. This is an empirical question, which can be answered by recording the response of binaural cells to the binaurally modulated stimuli used in the present study, allowing a direct comparison between ILD- and ITD-invoked decorrelation.

Interestingly, the question of "how compressed" the monaural inputs to the binaural processor are relates to another major problem in our understanding of binaural processing. An undesired feature of simple coincidence detectors (and of their continuous counterpart, the unnormalized crosscorrelation) is their sensitivity to diotic intensity fluctuations or variations. The trial-to-trial variations in intensity of the noise tokens will easily swamp the detector's

sensitivity to binaural disparity. This problem is discussed in depth by Van de Par et al. (2001) and Colburn and Isabelle (2001). A possible solution is to normalize or compress the monaural inputs, that is, to render them less sensitive to monaural intensity fluctuations. Such monaural compression will have the obvious side effect of also reducing the sensitivity to ILDs. From the analysis of Van de Par et al. (2001), it is clear that a strongly compressive transformation is needed to counteract the confounding effect of sample intensity; note that we also needed a highly compressive transformation to account for the ITD dominance in our data (Fig. 7B).

Whether compression of the monaural inputs followed by unnormalized crosscorrelation is a realistic description of binaural processing is again an empirical question. To find out whether the monaural inputs show such effective compression, one may measure the response of monaural cells to the stimuli used in the present study, followed by a cross-coincidence analysis (Joris 2003; Louage et al. 2004). It was shown by Louage et al. (2006) that a cross-coincidence analysis of single bushy cell responses provides the acuity needed to reproduce an exquisite sensitivity to interaural correlation changes (comparable to psychoacoustical thresholds). This indicates that the normalization problem raised by Van de Par et al. (2001) is already essentially solved at the level of the monaural inputs to the binaural processor—without the need for excitation-inhibition or "subtraction" schemes proposed by those authors. On the other hand, it is not known whether the effective normalization is realized in a way that also causes reduced ILD sensitivity. Incidentally, it is entirely possible that the monaural inputs do *not* show such reduced ILD sensitivity, whereas binaural cells at the level of the IC do. That would indicate a compression or normalization mechanism operating at the binaural level, e.g., through inhibitory inputs to the medial superior olive or the IC. Physiological experiments as suggested above may help clarify these issues.

Compression in cochlear mechanics

It seems natural to explain the reduced sensitivity to ILDs by the compression observed in cochlear-mechanical measurements (reviewed in Cooper 2006). Such a role of cochlear nonlinearity in binaural detection was proposed by Van de Par and Kohlrausch (1998) based on their narrowband, 4-kHz data, following analogous interpretations of monaural psychoacoustical data by Oxenham and Plack (1997). From Figure 7B, it would seem that our data can indeed be (partially) interpreted in terms of compressive cochlear growth functions. We do not favor such an explanation of our data.

It is unlikely that our wideband, low-frequency stimuli are subject to much cochlear compression at all. Reports of strong compression in cochlear mechanics are exclusively based on basilar membrane responses to single, high-frequency tones in the basal and mid turns of the cochlea. Such data are of little relevance to the stimuli used in low-frequency MLD studies. There exists evidence that the low-frequency mechanics of the apex is (much) less compressive than the high-frequency mechanics in the basal turns (Cooper 2006; Robles and Ruggero 2001). More importantly, wideband stationary stimuli like noise are known to greatly linearize the cochlear-mechanical reponse (De Boer and Nutall 1997, 2000). The same basal sites that yield highly compressive responses to single tones produce a perfectly Rayleigh-distributed envelope when stimulated with wideband noise (Recio-Spinoso et al. 2009). Thus for wideband noise stimuli, there is no envelope compression on the basilar membrane.

If cochlear compression is ruled out, it is important to identify alternative physiological correlates of compression of the monaural inputs. Apart from the neural transformation from the auditory nerve to the bushy cells discussed above, there are many nonlinear processes in the auditory periphery that may contribute to the effective compression of the monaural inputs: Boltzmann statistics of the transduction channels, nonlinear variation of receptor potential with hair bundle deflection, nonlinearities of synaptic release and postsynaptic currents, adaptation, all-or-nothing character of action potentials, etc. Note also that a full-fledged model of binaural detection eventually depends on the statistics of the neural correlates of the binaural stimulus parameters. Envelope compression by itself does not necessarily reduce the sensitivity to envelope fluctuations; it is the *accuracy* with which the envelope is "coded" that limits its role in performance.

Potential applications of binaural modulation

We have introduced a binaural modulation technique for the selective jittering of dynamic ITDs and/or ILDs in a binaural stimulus and applied it to a classic NoSπ condition with a 500-Hz signal and a wideband masker. The results were clear-cut: Interaural correlation of the stimulus waveform is a poor predictor of performance; dynamic ITDs are more important to the task than dynamic ILDs. Without further experiments, these results may not be generalized to other binaural listening tasks. Fortunately, the generic character of the modulation technique allows it to be applied to any binaural listening task. In the domain of decorrelation detection, the use of binaural modulation may provide an alternative to the frozen-noise approach of Goupell and Hartmann (2006, 2007a, b). Binaural modulation may also help

clarify the origins of the interesting effects of band-width and frequency on correlation detection (Gabriel and Colburn 1981; Culling et al. 2001). In the domain of binaural detection, obvious parameters to be varied are noise bandwidth and signal frequency. At first glance, one might expect that ITDs will cease to play a dominant role at higher (>1,500 Hz) signal frequencies due to the lack of high-frequency phase sensitivity (Johnson 1980). Detection at high frequencies, however, may well be dominated by *envelope* ITDs (or ILDs). Thus the question of the relative contributions of ILDs and ITDs is equally valid in the domain of high-frequency MLDs. Because our modulation technique is applied to the entire stimulus of a binaural listening task (fine structure and envelope alike), it is equally well suited for high-frequency conditions. Finally, binaural modulation may be used, not as a means of degrading the stimulus of a given binaural task but as the feature to be detected. With proper reference conditions, the absence of monaural cues in the stimuli allows one to directly test the sensitivity of the binaural system to dynamic fluctuations in ILDs and ITDs for a wide variety of stimuli, both in psychoacoustical and physiological experiments.

ACKNOWLEDGMENTS

Open Access

APPENDIX: COMPUTATIONAL DETAILS

Mixed modulation.

Amplitude modulation of a pure tone $x(t) = \cos \omega_c t$ may be expressed in two mathematically equivalent ways. The first is multiplication by a modulator:

$$x_{AM}(t) = (1 + m \cos \omega_m t) \cos \omega_c t. \qquad (1)$$

The second is by the addition of two sidebands

$$x_{AM}(t) = \cos \omega_c t + m/2[\cos (\omega_c - \omega_m)t + \cos (\omega_c + \omega_m)t]. \qquad (2)$$

The two expressions are mathematically equivalent, but the first one is more practical when generalizing from the modulation of pure tones to the modulation of arbitrary waveforms $x(t)$:

$$x_{AM}(t) = (1 + m \cos \omega_m t)x(t). \qquad (3)$$

A similar generalization of Eq. 2 is more cumbersome in that it requires adding sidebands to each individual component of the Fourier spectrum of $x(t)$.

The definition of QFM is generally stated in terms of added sidebands:

$$x_{QFM}(t) = \cos \omega_c t + m/2[\sin (\omega_c - \omega_m)t + \sin(\omega_c + \omega_m)t]. \qquad (4)$$

This differs from Eq. 2 only by a change of the phase of the sidebands. In order to find the multiplicative expression for QFM (the analog of Eq. 1), it is necessary to change to the complex-analytic version $z(t)$ of the waveforms by adding an imaginary part to $x(t)$ equal to its Hilbert transform $H[x(t)]$. A tone $x(t) = \cos \omega t$ thus becomes

$$z(t) = e^{i\omega t} \qquad (5)$$

and a complex-analytic Gaussian noise buffer is readily computed by starting in the complex Fourier domain while restricting the spectrum to only positive frequencies. The true waveform is always recovered by taking the real part of the complex-analytic waveform:

$$x(t) = Re[z(t)]. \qquad (6)$$

Imposing AM on a complex-analytic waveform $z(t)$ is described by the same multiplier of Eq. 3

$$z_{AM}(t) = (1 + m \cos \omega_m t)z(t). \qquad (7)$$

This time, however, QFM is described by a similar multiplication

$$z_{QFM}(t) = (1 - im \cos \omega_m t)z(t), \qquad (8)$$

which reduces to Eq. 4 after substitution of Eq. 5 and application of Eq. 6. Unlike Eq. 4, however, which only describes QFM of pure tones, Eq. 8 is applicable to arbitrary complex-analytic waveforms, thereby achieving the desired reduction of QFM to a multiplication. Equations 7 and 8 are special cases of the more general expression

$$z_\theta(t) = (1 + e^{i\theta} m \cos \omega_m t)z(t). \qquad (9)$$

Here the "modulation angle" θ determines the relative contributions of AM and QFM to the modulation; $\theta = 0$ and $\theta = \pi$ correspond to pure AM; $\theta = \pm\pi/2$ corresponds to pure QFM. The mixed modulation used in this study is obtained by choosing θ from the set of values $\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$; in Figure 2, these angles show up as the tilt of the dashed lines in Fig. 3 with respect to the vertical.

Binaural modulation.

The different types of binaural modulation summarized in Table 1 (main text) were realized by taking different combinations of θ for the two ears. Denoting the modulation angles for the left and right ear by θ_L and θ_R, respectively, and using the subscript notation for the binaural modulation types, these combination are:

$$\left(\theta_L, \theta_R\right)_d = (\pi/4, \pi/4) \qquad \text{(diotic)} \qquad (10a)$$

$$\left(\theta_L, \theta_R\right)_{mx} = (\pi/4, 5\pi/4) \qquad \text{(binaural mixed)} \quad (10b)$$

$$\left(\theta_L, \theta_R\right)_a = (\pi/4, 3\pi/4) \qquad \text{(binaural AM)} \qquad (10c)$$

$$\left(\theta_L, \theta_R\right)_\phi = (\pi/4, 7\pi/4). \qquad \text{(binaural QFM).} \quad (10d)$$

Modulation was always applied to the whole waveform, i.e., both masker and signal (if present). Thus, denoting the unmodulated complex waveforms to the two ears by $z_L(t)$ and $z_R(t)$, the modulated waveforms were obtained by a straightforward application of Eq. 9

$$z_{L,mod}(t) = \left(1 + e^{i\theta_L} m \cos \omega_m t\right) z_L(t) \qquad (11a)$$
$$z_{R,mod}(t) = \left(1 + e^{i\theta_R} m \cos \omega_m t\right) z_R(t). \qquad (11b)$$

For very large modulation depths ($m \cong 1$), the modulator terms in parentheses acquire a nonzero mean phase for some values of θ. To prevent the binaural modulation from evoking a nonzero static interaural phase difference in any of our binaural configurations, a static phase factor was always applied to the modulator terms in order to render their time average real-valued.

Next consider the effect of the modulation of Eqs. 11a and 11b on the interaural correlation of the stimulus. The normalized crosscorrelation of two real waveforms $x_{1,2}(t) = \text{Re}[z_{1,2}(t)]$ is given by

$$\rho = \langle x_1 x_2 \rangle \Big/ \sqrt{\langle x_1^2 \rangle \langle x_2^2 \rangle} = \left\langle \text{Re}\left[z_1 z_2^*\right] \right\rangle \Big/ \sqrt{\left\langle z_1 z_1^* \right\rangle \left\langle z_2 z_2^* \right\rangle}, \qquad (12)$$

where $\langle \rangle$ denotes the expected value, z^* is the complex conjugate of z, and where we used $\langle z_1 \rangle = 0 = \langle z_1 z_2 \rangle$, etc., because the analytical waveforms contain only positive frequencies. Substitution of Eqs. 11a and 11b yields

$$\rho_m = \left\langle \text{Re}\left[z_L z_R^*\right]\left(1 + \tfrac{m^2}{2}\cos \Delta\right) - \text{Im}\left[z_L z_R^*\right]\tfrac{m^2}{2}\sin \Delta \right\rangle \Big/ \left(\left(1 + \tfrac{m^2}{2}\right)\sqrt{\left\langle z_L z_L^* \right\rangle \left\langle z_R z_R^* \right\rangle}\right), \qquad (13)$$

where $\Delta = \theta_L - \theta_R$, and it is assumed that z_L and z_R do not contain components at the modulation frequency ω_m and its second harmonic $2\omega_m$. For the stimuli used in this study, the unmodulated waveforms z_L and z_R are either identical (noise-alone or reference interval) or the sum of identical noise and an antiphasic tone (signal or target interval). In neither case is there a correlation between the stimulus in one ear and the Hilbert transform ($\pi/2$ phase-shifted version) of the stimulus in the other ear. Thus $\left\langle \text{Im}\left[z_1 z_2^*\right] \right\rangle$ vanishes, and Eq. 13 reduces to

$$\rho_m = \frac{1 + \tfrac{m^2}{2}\cos \Delta}{\left(1 + \tfrac{m^2}{2}\right)} \left\langle \text{Re}\left[z_L z_R^*\right]\right\rangle \Big/ \left(\sqrt{\left\langle z_L z_L^* \right\rangle \left\langle z_R z_R^* \right\rangle}\right),$$

$$= \frac{2 + m^2 \cos \Delta}{2 + m^2} \rho_0, \qquad (14)$$

where ρ_0 is the correlation of the unmodulated stimulus. Thus the effect of modulation on the interaural correlation is a reduction of ρ by a factor $\beta = (2 + m^2 \cos \Delta)/(2 + m^2)$. For the four different types of binaural modulation listed in Eqs. 10a and 10b, this reduction factor β equals

$$\beta_d = 1 \qquad \text{(diotic)} \qquad (15a)$$

$$\beta_{mx} = \frac{2 - m^2}{2 + m^2} \qquad \text{(binaural mixed)} \quad (15b)$$

$$\beta_a = \frac{2}{2 + m^2} \qquad \text{(binaural AM)} \qquad (15c)$$

$$\beta_\phi = \frac{2}{2 + m^2} \qquad \text{(binaural QFM)} \quad (15d)$$

As anticipated, binaural AM and binaural QFM have exactly the same effect on the interaural correlation of the stimulus. Any correlation-based model of binaural detection will thus predict the same effect of these two types of binaural modulation.

Relation to NρSπ thresholds.

The problem is to relate the modulated masking conditions $(N0S\pi)_a$, $(N0S\pi)_\phi$, and $(N0S\pi)_{mx}$ to the masking conditions NρSπ, in which the noise correlation is varied by mixing independent noise waveforms. It is important to realize that the reduction of interaural correlation described by the Eqs. 15a–15d applies to the entire stimulus. Not only is the correlation of the noise masker reduced but also the correlation of the noise-plus-signal combination is reduced by the same factor. In the context of models based on interaural correlation, the former effect is equivalent to using a decorrelated noise masker; the latter effect is equivalent to using a signal whose effect on the correlation of the stimulus is less than that of a Sπ signal (stated differently, the antiphasic portion of the signal is reduced, thereby reducing its ability to evoke binaural disparity). In short, the masker is decorrelated, and the effective power of the signal is reduced.

For a given reduction factor β from Eqs. 15a–15d, the noise correlation is $\rho=\beta$. The reduced effective power of the signal can be evaluated by applying Eq. 12. The effect is not uniform, in that it depends on the S/N ratio, but for small S/N ratios, it is well approximated by a uniform reduction of effective signal power by a factor $2\beta/(\beta+1)$. This leads to the following prescription to predict the "modulated thresholds" from NρSπ data:

1. For a given binaural modulation type and depth m, determine the reduction factor β from Eqs. 15a–15d.
2. Select the masked threshold in decibels from the NρSπ condition with $\rho=\beta$.
3. Correct for the reduced effective signal power by adding $10\log(\beta+1)/2\beta$ dB.
4. If this leads to a negative MLD, use the monaural threshold instead.

This prescription was used to compile the predictions of Figure 7. For a more elaborate description of this type of modeling, see Van der Heijden and Trahiotis 1999. The NρSπ thresholds needed in step 2 were estimated from the measured N0Sπ and N0S0 thresholds using Eq. 2 of Van der Heijden and Trahiotis (1997).

REFERENCES

BLAUERT J (1981) Lateralization of jittered tones. J Acoust Soc Am 70:694–698

BOEHNKE SE, HALL SE, MARQUARDT T (2002) Detection of static and dynamic changes in interaural correlation. J Acoust Soc Am 112:1617–1626

BUUS S (1985) Release from masking caused by envelope fluctuations. J Acoust Soc Am 78:1958–1965

CHERRY C, SAYERS BMA (1959) On the mechanism of binaural fusion. J Acoust Soc Am 31:535

COLBURN HS (1973) Theory of binaural interaction based on auditory-nerve data. I. General strategy and preliminary results on interaural discrimination. J Acoust Soc Am 54:1458–1470

COLBURN HS, DURLACH NI (1978) Models of binaural interaction. In: Carterette (ed) Handbook of perception, Vol 4 Hearing. Academic, New York, pp 467–518

COLBURN HS, ISABELLE SK (2001) Physiologically based models of binaural detection. In: Breebaart DJ, Houtsma AJM, Kohlrausch A, Prijs V, Schoonhoven R (eds) Physiological and psychophysical bases of auditory function. Shaker, Maastricht, pp 161–168

COOPER NP (2006) Compression in the peripheral auditory system. In: Bacon SP, Fay RR, Popper AN (eds) Compression: from cochlea to cochlear implants. New York, Springer, pp 18–61

CULLING JF, COLBURN HS, SPURCHISE M (2001) Interaural correlation sensitivity. J Acoust Soc Am 110:1020–1029

DE BOER E, NUTALL AL (1997) The mechanical waveform of the basilar membrane. I. Frequency modulations ("glides") in impulse responses and cross-correlation functions. J Acoust Soc Am 101:3583–3592

DE BOER E, NUTALL AL (2000) The mechanical waveform of the basilar membrane. III. Intensity effects. J Acoust Soc Am 107:1497–1507

DOLAN RE, ROBINSON DE (1967) Explanation of masking-level differences that result from interaural intensive disparities of noise. J Acoust Soc Am 42:977–981

DOMNITZ RH, COLBURN HS (1976) Analysis of binaural detection models for dependence on interaural target parameters. J Acoust Soc Am 59:598–601

DURLACH NI (1964) Note on binaural masking-level differences at high frequencies. J Acoust Soc Am 36:576–581

DURLACH NI (1972) Binaural signal detection: equalization and cancellation theory. In: Tobias JV (ed) Foundations of modern auditory theory, vol II. Academic, New York, pp 365–466

GABRIEL KJ, COLBURN HS (1981) Interaural correlation discrimination: I. Bandwidth and level dependence. J Acoust Soc Am 69:1394–1401

GOUPELL MJ, HARTMANN WM (2006) Interaural fluctuations and the detection of interaural incoherence: bandwidth effects. J Acoust Soc Am 119:3971–3986

GOUPELL MJ, HARTMANN WM (2007a) Interaural fluctuations and the detection of interaural incoherence. II. Brief duration noises. J Acoust Soc Am 121:2127–2136

GOUPELL MJ, HARTMANN WM (2007b) Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models. J Acoust Soc Am 122:1029–1045

GRANTHAM DW, WIGHTMAN FL (1978) Detectability of varying interaural temporal differences. J Acoust Soc Am 63:511–523

HAFTER ER, CARRIER SC (1970) Masking level differences obtained with pulsed tonal maskers. J Acoust Soc Am 47:10141–11047

HIRSH IJ (1948) The influence of interaural phase on interaural summation and inhibition. J Acoust Soc Am 20:536–544

JEFFRESS LA, BLODGETT HC, SANDEL TT, WOOD CL (1956) Masking of tonal signals. J Acoust Soc Am 28:416–426

JOHNSON DH (1980) The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. J Acoust Soc Am 68:1115–1122

JORIS PX (2003) Interaural time sensitivity dominated by cochlea-induced envelope patterns. J Neurosci 23:6345–63450

BERNSTEIN LR, VAN DE PAR S, TRAHIOTIS C (1999) The normalized correlation: accounting for NoSπ thresholds obtained with Gaussian and "low-noise" masking noise. J Acoust Soc Am 106:870–876

BERNSTEIN LR, TRAHIOTIS C, AKEROYD MA, HARTUNG K (2001) Sensitivity to brief changes of interaural time and interaural intensity. J Acoust Soc Am 109:1604–1615

JORIS PX, YIN TC (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am 91:215–232

JORIS PX, YIN TC (1998) Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. J Neurophysiol 79:253–269

JORIS PX, CARNEY LH, SMITH PH, YIN TC (1994) Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. J Neurophysiol 71:1022–1036

JORIS PX, VAN DE SANDE B, RECIO-SPINOSO A, VAN DER HEIJDEN M (2006) Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. J Neurosci 26:279–289

LOUAGE DHG, VAN DER HEIJDEN M, JORIS PX (2004) Temporal properties of responses to broadband noise in the auditory nerve. J Neurophysiol 91:2051–2065

LOUAGE DHG, JORIS PX, VAN DER HEIJDEN M (2006) Decorrelation sensitivity of auditory nerve and anteroventral cochlear nucleus fibers to broadband and narrowband noise. J Neurosci 26:96–108

MCFADDEN D, JEFFRESS LA, ERMEY HL (1971) Differences of interaural phase and level in detection and lateralization: 250 Hz. J Acoust Soc Am 50:1484–1493

OSMAN E (1971) A correlation model of binaural masking level differences. J Acoust Soc Am 50:1494–1511

OXENHAM AJ, PLACK CJ (1997) A behavioral measure of basilarmembrane nonlinearity in listeners with normal and impaired hearing. J Acoust Soc Am 101:3666–3675

RECIO-SPINOSO A, NARAYAN SS, RUGGERO MA (2009) Basilar-membrane responses to noise at a basal site of the chinchilla cochlea: quasilinear filtering. J Assoc Res Otolaryngol. doi:10.1007/s10162-009-0172-0)

RICHARDS VM (1992) The detectability of a tone added to narrow bands of equal-energy noise. J Acoust Soc Am 91:3424–3435

ROBINSON DE, JEFFRESS LA (1963) Effect of varying the interaural noise correlation on the detectability of tone signals. J Acoust Soc Am 35:1947–1952

ROBLES L, RUGGERO MA (2001) Mechanics of the mammalian cochlea. Physiol Rev 81:1305–1352

SIVEKE I, EWERT SD, GROTHE B, WIEGREBE L (2008) Psychophysical and physiological evidence for fast binaural processing. J Neurosci 28:2043–2045

STELLMACK MA, VIEMEISTER NF, BYRNE AJ (2005) Monaural and interaural temporal modulation transfer functions measured with 5-kHz carriers. J Acoust Soc Am 118:2507–2518

THOMPSON ER, DAU T (2008) Binaural processing of modulated interaural level differences. J Acoust Soc Am 123:1017–1029

VAN DE PAR S, KOHLRAUSCH A (1998) Diotic and dichotic detection using multiplied-noise maskers. J Acoust Soc Am 103:2100–2110

VAN DE PAR S, TRAHIOTIS C, BERNSTEIN LR (2001) A consideration of the normalization that is typically included in correlation-based models of binaural detection. J Acoust Soc Am 109:830–833

VAN DER HEIJDEN M, KOHLRAUSCH A (1995) The role of envelope fluctuations in spectral masking. J Acoust Soc Am 97:1800–1807

VAN DER HEIJDEN M, TRAHIOTIS C (1997) A new way to account for binaural detection as a function of interaural noise correlation. J Acoust Soc Am 101:1019–1022

VAN DER HEIJDEN M, TRAHIOTIS C (1998) Binaural detection as a function of interaural correlation and bandwidth of masking noise: Implications for estimates of spectral resolution. J Acoust Soc Am 103:1609–1614

VAN DER HEIJDEN M, TRAHIOTIS C (1999) Masking with interaurally delayed stimuli: the use of "internal" delays in binaural detection. J Acoust Soc Am 105:388–399

WAGNER H (1991) A temporal window for lateralization of interaural time difference by barn owls. J Comp Physiol [A] 169:281–289

WAGNER H (1992) On the ability of neurons in the barn owl's inferior colliculus to sense brief appearances of interaural time difference. J Comp Physiol [A] 170:3–11

WEBSTER FA (1951) The influence of interaural phase on masked thresholds I. the role of interaural time-deviation. J Acoust Soc Am 23:452–462

ZUREK PM (1991) Probability distributions of interaural phase and level differences in binaural detection stimuli. J Acoust Soc Am 90:1927–1932

ZUREK PM, DURLACH NI (1987) Masker-bandwidth dependence in homophasic and antiphasic tone detection. J Acoust Soc Am 81:459–464