

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/30435/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Culling, John Francis 2011. Subcomponent cues in binaural unmasking. *Journal of the Acoustical Society of America* 129 (6) , pp. 3846-3855. 10.1121/1.3560944 file

Publishers page: <http://dx.doi.org/10.1121/1.3560944>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



# Subcomponent cues in binaural unmasking<sup>a)</sup>

John F. Culling<sup>b)</sup>

School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, United Kingdom

(Received 1 October 2010; revised 28 January 2011; accepted 10 February 2011)

The addition of a signal in the NOS $\pi$  binaural configuration gives rise to fluctuations in interaural phase and amplitude. Sensitivity to these individual cues was measured by applying sinusoidal amplitude modulation (AM) or quasi-frequency modulation (QFM) to a band of noise. Discrimination between interaurally in-phase and out-of-phase modulation was measured using an adaptive task for narrow bands of noise at center frequencies from 250 to 1500 Hz, for modulation rates of 2–40 Hz, and with or without flanking bands of diotic noise. Discrimination thresholds increased steeply for QFM with increasing center frequency, but increased only modestly for AM, and mainly for modulation rates below 10 Hz. Flanking bands of noise increased thresholds for AM, but had no consistent effect for QFM. The results suggest that two underlying mechanisms may support binaural unmasking: one most sensitive to interaural amplitude modulations that is susceptible to across-frequency interference, and a second, most sensitive to interaural phase modulations that is immune to such effects.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3560944]

PACS number(s): 43.66.Pn, 43.66.Dc, 43.66.Ba [RYL]

Pages: 3846–3855

## I. INTRODUCTION

A low-frequency tonal signal is more easily detected in noise when the noise is presented interaurally in-phase and the signal is interaurally out-of-phase (NOS $\pi$ ) than when both are in phase (NOS0). This “binaural unmasking” indicates that the brain is sensitive to differences in the waveform of the sound at the two ears and uses them in order to detect the presence of the signal (Hirsh, 1948). The mechanism(s) underlying this phenomenon remain a matter of controversy, arguably because different proposed cues to binaural detection are highly correlated across different types of stimulus, making predictions based on each prospective cue very similar (Colburn and Durlach, 1978, p. 509). The present study separately controls two potential cues: Interaural time delays (ITDs) and interaural level differences (ILDs) in listening configurations designed to mimic binaural unmasking paradigms. The results are offered in an effort to provide a more varied source of comparison data for models of binaural unmasking.

Previous studies have attempted to separate the roles of ILDs and ITDs in binaural unmasking by several different means. Most, like the present study, have experimentally controlled the presence of ILDs and ITDs in an attempt to measure their individual effects. Other studies have measured variations in detection for reproducible-noise stimuli and attempted to interpret them in terms of the ILDs and ITDs that occur in those individual stimuli. One study has added additional modulation in ILD or ITD to the entire stimulus in order to selectively jam one cue or the other.

## A. Experimental control studies

Hafer and Carrier (1970), McFadden *et al.* (1971), and Yost (1972) measured the sensitivity to static ILDs and ITDs by adding tonal signals to tonal maskers or noise signals to identical noise maskers using controlled phase relations. The results of such studies can be difficult to interpret, because static ILDs and ITDs give rise to a perception of laterality; the overall sound is heard in a different location within the head. While “position variable” theories of binaural unmasking explicitly associate laterality with unmasking as part of the same process (Stern and Colburn, 1985), others, such as EC theory (Durlach, 1963, 1972; Culling and Summerfield, 1995; Culling, 2007), regard the two as independent. Moreover, in the typical configuration of a tonal signal and a random noise masker, the presence of the signal causes fluctuating ITDs and ILDs. Consequently, van de Par and Kohlrausch (1998) developed stimuli based on a multiplied noise technique in which fluctuating ILDs or ITDs were selectively produced from a tone-and-noise stimulus. While their method created a more ecologically valid stimulus, it did not entirely eliminate cues from laterality, because the mean ILD or ITD could be non-zero. Nonetheless, van de Par and Kohlrausch’s data showed that sensitivity to fluctuating ITDs declined with increasing center frequency, whereas sensitivity to fluctuating ILDs, while poorer at low frequencies, was relatively frequency independent. Figure 1 reproduces their data, averaged across listeners. Although these patterns differed, van de Par and Kohlrausch suggested that peripheral non-linearities, rather than separate central processors, could explain the differences. Specifically, peripheral compression impairs the encoding of amplitude modulations at all carrier frequencies, explaining why thresholds for fluctuating ILDs were higher than those for fluctuating ITDs at low frequencies. Meanwhile, loss of phase locking at higher frequencies, progressively impairs the encoding of fine-structure ITDs as frequency increases, explaining the increase in thresholds with frequency for fluctuating ITDs.

<sup>a)</sup>Portions of this work were presented at the International Symposium on Hearing in Salamanca, Spain, and appeared in *The Neurophysiological Bases of Auditory Perception* (Springer, New York, 2010).

<sup>b)</sup>Author to whom correspondence should be addressed. Electronic mail: CullingJ@cf.ac.uk

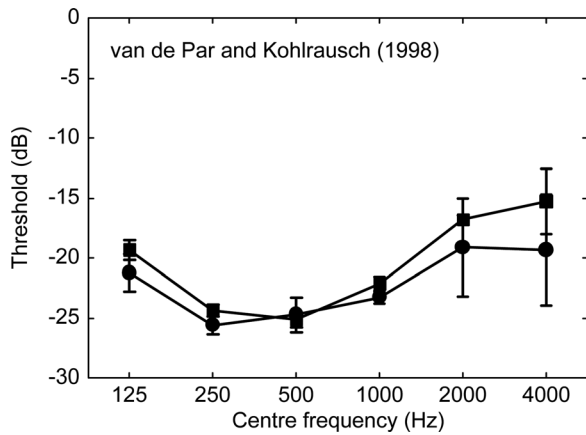


FIG. 1. Data from experiment 2 of van de Par and Kohlrausch (1998). Mean thresholds (four listeners) for detection of a tonal signal in a 25-Hz wide band of noise, constructed such that the presence of the signal created either interaural phase modulations (squares) or interaural amplitude modulations (circles). Error bars are one standard error of the mean ( $n = 4$ ).

## B. Reproducible noise studies

A number of proposed models/theories have identified interaural correlation as a potential decision variable in binaural unmasking (Osman, 1971; Colburn, 1973, 1977; Durlach *et al.* 1986; Bernstein and Trahiotis, 2003). Interaural correlation is reduced at the signal frequency compared to adjacent frequency channels or to a diotic comparison stimulus, so listeners may detect this reduction in interaural correlation and interpret it as evidence of a signal.

Goupell and Hartmann (2006) examined this idea by measuring listeners' ability to discriminate individual frozen noises (narrow bands centered on 500 Hz), whose interaural correlations were all equal (0.9922) from similar diotic noises. Listeners' ability to make this discrimination varied markedly across different noise samples, particularly for narrowband stimuli, suggesting that the fixed cue of interaural correlation was not the parameter that controlled detection. Goupell and Hartmann (2007) tested a variety of models to account for these data, including ones that extracted the mean ILD and/or ITD, the fluctuation in ILD and/or ITD, and the interaural correlation. They found that models that processed fluctuations in ILDs and ITDs independently, and then combined the information from each, performed the best. Goupell and Hartmann's work thus suggests that binaural unmasking may be produced by more than one mechanism operating in parallel.

A problem that Goupell and Hartmann encountered in their modeling was that extraction of the time-varying ILD and ITD from the stimuli (using the Hilbert transform) resulted in occasionally extreme values of ITD, even though the interaural correlation was very high. These extreme values occurred at moments in time when there was little energy in the stimulus, but could nonetheless have a large influence on the measured ITD variance. Goupell and Hartmann ameliorated the influence of this effect by placing a threshold limit on the size of ITD to be including in the input to each of their models.

Davidson *et al.* (2009) also employed the Hilbert transform in order to separate ILDs and ITDs, but they employed

it in the preparation of their stimuli. The envelopes and fine structures of narrowband NOS $\pi$  stimuli (centered on 500 Hz) were extracted using the Hilbert transform and then "chimeric" stimuli were created by combining the envelope of one stimulus with the fine structure of another. Thus, a pair of NOS $\pi$  stimuli could be transformed into a pair of chimeric NOS $\pi$  stimuli, whose ILDs and ITDs were swapped. As in Goupell and Hartmann's modeling, because large changes in ITD can occur at points of low energy in the stimulus, the chimeric stimuli could contain anomalies; when a large change in ITD was recombined with a temporal envelope which does not have a dip in energy at that point in time, the result was wideband spectral splatter. Davidson *et al.* minimized the influence of this effect by selecting stimuli whose bandwidth was relatively unaltered by the processing. Measurements of tone detection among the original reproducible noises and their chimeric combinations showed that the successful detection of tones in individual stimuli was associated more strongly with the envelopes of those stimuli than with their fine structures, but that both were involved. Moreover it appeared that they were *not* processed independently.

## C. ITD/ILD masking

Van der Heijden and Joris (2010) adopted a different approach. Rather than experimentally manipulating particular cues to the presence of a tonal signal or correlating performance with the prominence of one cue or another in individual stimuli, they selectively interfered with one cue or the other. This interference was achieved by creating NOS0 or NOS $\pi$  stimuli and then imposing additional modulation of ILD or interaural phase difference (IPD) on the entire stimulus (tone and noise). These additional modulations were added to the stimuli in the complex analytic domain in the form of interaural amplitude modulation (AM) and/or quasi-frequency modulation (QFM). Both modulations were added to both ears, but one was interaurally in-phase and the other was interaurally out-of-phase. Interaurally out-of-phase AM generates additional ILD modulation which is unrelated to the presence of the tonal signal and therefore masks the information from that cue. Similarly out-of-phase QFM generates additional ITD modulation and masks information from that cue. Thus, if the cue that listeners use is jammed in this way, then detection thresholds should be elevated.

An advantage of this scheme is that it can be readily applied to tones in broadband as well as narrowband noise. Van der Heijden and Joris used a 500-Hz tone in a 100 to 3000-Hz band of Gaussian noise. They found that thresholds were elevated more by ITD modulation than by ILD modulation, indicating that for 500-Hz tones in broadband noise modulation of ITD is the dominant cue in NOS $\pi$  detection.

## D. The present experiment

The present experiment is similar to that of van de Par and Kohlrausch (1998), but with two innovations. First, van der Heijden and Joris' technique was used to introduce ILDs and ITDs. Rather than add signals to maskers, the ILD and ITD modulations were generated directly by manipulating a band of noise, simulating the presence of a signal. This

method precisely controls these fluctuating values so that they can have a specified frequency and, provided an integer number of cycles is presented, zero mean. The advantage of keeping the mean at zero is that these stimuli should produce no net percept of laterality.

Fluctuating ILDs were produced by applying interaurally out-of-phase AM to the noise band. AM changes the long term spectrum of a sound, such that each spectral component is divided into three. In order to avoid introducing a confounding spectral cue, listeners were required to discriminate interaurally out-of-phase modulation from interaurally in-phase modulation. Fluctuating ITDs were produced by applying interaurally out-of-phase QFM, which was similarly in-phase or out-of-phase across the ears. QFM is not pure frequency modulation (FM); it includes a small component of AM. However, like AM, QFM divides each component into three, whereas pure FM produces an extensive series of side components. QFM is thus more spectrally compact. This feature was advantageous for the second innovation.

Binaural unmasking experiments typically use either narrow or broadband noises as maskers. Unmasking effects differ substantially between these two types of masker. In particular, the size of the unmasking effect, the binaural masking level difference (BMLD), is highly dependent on frequency for a broadband masker (Hirsh and Burgeat, 1958), but is largely independent of frequency for a narrowband masker (McFadden and Passanen, 1974). Moreover, BMLDs are very consistent across listeners for a broadband masker but very variable across listeners for a narrowband masker (Koehnke *et al.* 1986; Bernstein *et al.* 1998). Since van de Par and Kohlrausch only used a narrowband masker, their results cannot be generalized to all forms of binaural unmasking. Meanwhile, van der Heijden and Joris only examined the broadband case. In the present experiment, flanking bands of diotic noise were added to the manipulated band of noise in order to simulate the broadband case. In order to avoid disrupting the modulation therein, a spectral notch was left on either side of the manipulated band, which reduced masking of the modulation sidebands by the flanking noise. The use of QFM, rather than FM, ensured that any residual masking would be the same as that in the AM case, because the modulation sidebands had identical spectra. It is also noteworthy that interaural correlation is related to modulation index,  $m$ , in the same way for both AM and QFM and that the resulting interaural correlation is less variable than in stimuli generated by mixing interaurally correlated and uncorrelated noise in predetermined ratios.

## II. MAIN EXPERIMENT

### A. Stimuli

The target bands were 1-ERB-wide (Moore and Glasberg, 1983) bands of Gaussian noise centered at 250, 500, 750, 1000, 1250, or 1500 Hz. Interaural modulations in amplitude and phase were generated by modulating these noise bands with different modulation phases at each ear. Following a Hilbert transform, AM and QFM were applied to the complex analytic signals by multiplying them by  $1 + m \cdot \sin$

$(\omega t + \varphi)$ , where  $t$  is time,  $\omega$  is the modulation rate (radians/s), and  $m$  is a modulation index which has a real value for AM and an imaginary value for QFM. For an interaurally modulated stimulus,  $\varphi$  was 0 for the left channel and  $\pi$  radians for the right channel; while for a comparison stimulus it was 0 for both channels. Real signals for presentation were then derived using inverse Hilbert transformation.

Where flanking bands were used, these were independently generated, of equal spectrum level to the target band and separated from it by 1-ERB-wide spectral notches. The flanking bands otherwise filled out the spectrum from 0 to 3 kHz.

All stimuli were 500 ms in duration including 10-ms raised-cosine onset/offset ramps. The stimuli were prepared digitally using MATLAB, a high-level computational programming language, at a sampling rate of 44.1 kHz and a 16-bit sample depth. Freshly generated noise was created for each trial. Filtering was first performed in the frequency domain by zeroing the amplitude of unwanted frequencies. Modulation was then applied to the target band. Flanking bands were added before the onset/offset ramps were finally applied.

Digital-to-analog conversion was performed using an Edirol UA20 soundcard (Roland, Shizuoka, Japan). The analog signals were amplified by a MTR HPA-2 headphone amplifier and presented through Sennheiser HD650 headphones (Wennbostel, Germany) in a single-walled IAC booth at 44 dB(A) for a stimulus with no flanking bands and 64 dB(A) for a stimulus with flanking bands.

### B. Procedure

Five listeners, including the author, took part in the experiment. During training, one listener proved unable to achieve comparable thresholds in the narrowband conditions to the other four and was dropped from the study. Once trained to asymptote, the remaining four listeners attended 30 1-h data-collection sessions. Each session tested one of the five modulation rates (2, 4, 10, 20, or 40 Hz) either with or without flanking bands. Each session was repeated three times. Within each session, twelve thresholds were measured, one for each center frequency (250, 500, 750, 1000, 1250, or 1500 Hz) and type of modulation (FM or QFM). Sessions using and not using flanking bands were alternated, but the different modulation rates were tested in a random sequence.

Threshold modulation index was measured using a 2-interval, 2-alternative, forced-choice task in a 2-down/1-up adaptive track (Levitt, 1971) with immediate trial-by-trial feedback. The adaptive track began at  $m = 1$ , and  $m$  was divided or multiplied by 1.2 for the first four reversals and then by 1.1 for the subsequent ten reversals. The value of  $m$  on the last ten reversals was averaged to give the measured threshold. No ceiling was applied to the value of  $m$ .

### C. Results

The mean results averaged across four listeners are shown in Fig. 2. Across all five modulation rates, there is a strong effect of center frequency with higher thresholds at

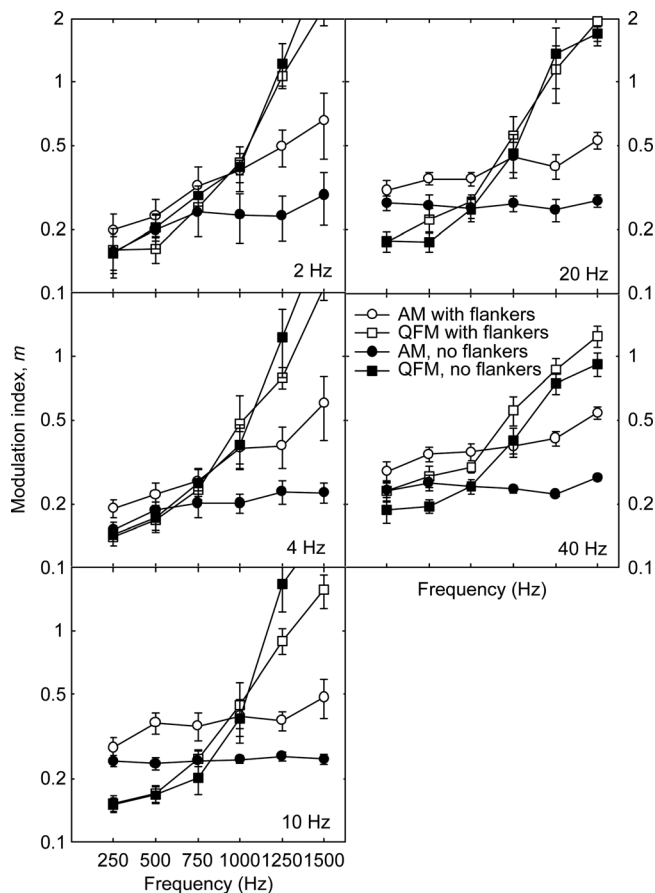


FIG. 2. Data from the main experiment. Each panel shows mean threshold modulation index as a function of center frequency at different modulation frequencies. Open symbols are data observed when flankers were included in the stimuli and closed symbols are data with the flankers absent. Squares are thresholds for interaural phase modulation and circles are for interaural amplitude modulation. Error bars are one standard error of the mean ( $n=4$ ).

higher center frequencies [ $F(5,15) = 67.5, p < 0.0001$ ]. However, this effect is quite modest for the AM condition and very strong for the QFM condition, producing a significant interaction [ $F(5,15) = 82.5, p < 0.0001$ ]. The relatively elevated thresholds for QFM at high frequencies also produced a significant main effect of modulation type, in which QFM thresholds were higher on average [ $F(1,3) = 34, p < 0.02$ ]. The presence of flanking bands of noise produced an elevation of thresholds in the AM condition alone, resulting in an interaction between their presence and the type of modulation [ $F(1,3) = 11.7, p < 0.05$ ]. This interaction was also responsible for an overall elevation of thresholds in the presence of flanking bands [ $F(1,3) = 10.8, p < 0.05$ ] and it also increased with center frequency, resulting in a three-way interaction [ $F(20,60) = 9.6, p < 0.0005$ ].

Looking across the five panels of Fig. 2, the effect of center frequency on QFM thresholds is greater at low modulation rates, resulting in interactions between center frequency and modulation rate [ $F(20,60) = 9.2, p < 0.0001$ ] and between that effect and modulation type [ $F(20,60) = 5.6, p < 0.0001$ ]. Finally, the effect of center frequency is greater for QFM when flankers are absent than present, but the reverse is true for AM, resulting in a center-frequency  $\times$  flanker-presence  $\times$  modulation-type interaction [ $F(5,15) = 9.6, p < 0.0005$ ].

## D. Discussion

### 1. Effects of modulation type and center frequency

Where flanking bands were absent, the data can be compared with those of van de Par and Kohlrausch (1998). Thresholds increased steeply with center frequency in the QFM condition, which was cued by modulation in interaural phase. In contrast, thresholds were roughly constant as a function of center frequency in the AM condition which was cued by interaural amplitude modulation. Across all of these conditions interaural correlation is related to  $m$  in the same way, so thresholds in terms of interaural correlation also varied widely. This pattern of results is similar to that observed in the second experiment of van de Par and Kohlrausch (1998) using a 25-Hz wide multiplied-noise masker. Van de Par and Kohlrausch also observed very little effect of center frequency on thresholds where only interaural amplitude modulations were present. They also observed a steep increase in thresholds with center frequency where only interaural phase modulations were present. As in the present data, thresholds were lower for interaural phase modulations at low frequencies, but lower for interaural amplitude modulations at higher frequencies. The cross-over occurred at about 500 Hz in their mean data. In the present experiment, using somewhat wider bands of noise, it occurred at about 750 Hz.

Van de Par and Kohlrausch attributed the rise with frequency in thresholds cued by interaural phase modulation to loss of phase locking, and the relatively high thresholds cued by interaural amplitude modulations at low frequencies to peripheral compression. In this way, they argued that both cues may be detected by a single mechanism with apparent differences in sensitivity being due to transformations of the stimuli during peripheral encoding. Thus, a correlation model could potentially explain the pattern of data, despite the fact that the interaural correlations of the threshold stimuli varied widely. Consequently, we need to draw a distinction between the external interaural correlation of the stimuli at the ears,  $\rho_e$ , and the internal interaural correlation following peripheral encoding  $\rho_i$ . In order to determine the value of  $\rho_i$ , a model of peripheral transduction is needed.

In order to account for lateralization data at high frequencies, Bernstein and Trahiotis (2003) developed a model of peripheral transduction that included the loss of fine temporal structure with increasing frequency and compression of the waveform envelope. The model incorporated a simulated auditory filter (Patterson *et al.*, 1995) centered on the stimulus frequency. The filtered waveform was then compressed using an exponent of 0.23, half-wave rectified, squared, and low-pass filtered using a fourth-order Butterworth filter with a corner frequency of 425 Hz. Such a model can be applied to the present data (Appendix). Figure 3 presents the empirical data from each modulation rate for the no-flanker conditions (symbols) against the predictions of such a model (black lines). The model has the same structure as that of Bernstein and Trahiotis (the predictions of which are shown by gray lines), but the parameters (low-pass filter order and cut-off, compression exponent and threshold for  $\rho_i$ ) are optimized for the present data set. The model

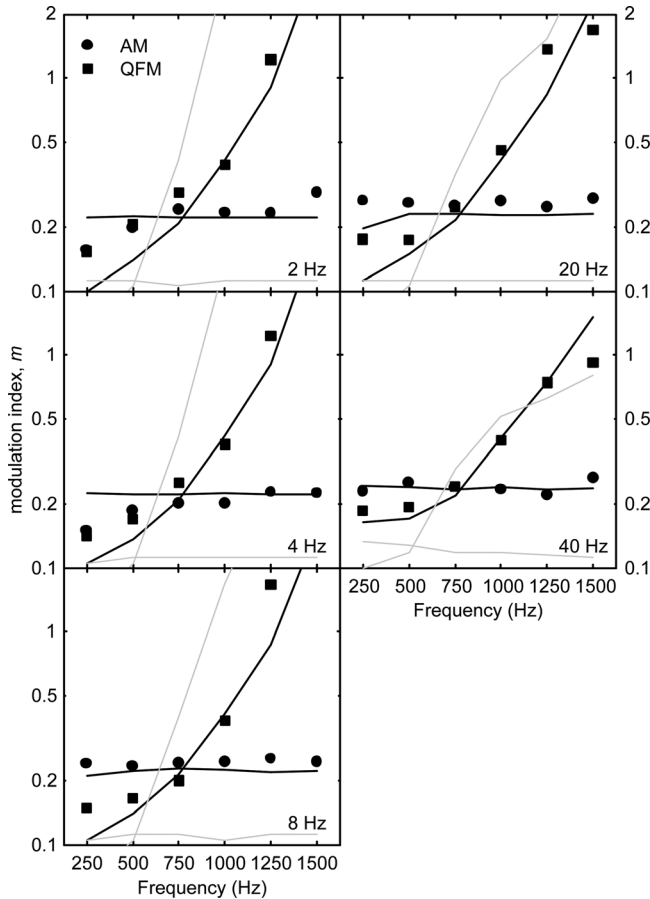


FIG. 3. Mean data from the conditions with no flankers (symbols) plotted against the predictions for the same stimuli of a model (black lines) based on that of [Bernstein and Trahiotis \(2003\)](#) using parameters fitted to the same data. For the data, means are averaged across four listeners. The predictions of the model using Bernstein and Trahiotis' low-pass filter and compression parameters are also shown (gray lines).

provides a good fit to the data with no flanking bands for a range of different parameter combinations. A compression ratio of 0.233 for consistency with [Bernstein and Trahiotis \(1996, 2003\)](#), but the slope of the increase in threshold with center frequency in the QFM condition was markedly shallower than would be predicted by a fourth-order filter at 425 Hz. The black lines in Fig. 3 show predictions from a third-order filter with 614-Hz cut-off, which produces a closely matching slope. Threshold  $\rho_i$  was 0.989. The model captures the principal trends in the data, particularly at the higher modulation rates, indicating that peripheral non-linearities offer a sufficient account of the data when flanking bands are absent.

The gray lines in Fig. 3 show equivalent results using the parameters published by [Bernstein and Trahiotis \(2003\)](#) in order to account for their lateralization data. The threshold value of  $\rho_i$  for this implementation of their model was the only parameter which was optimized in order to fit the present data. The optimal value was found to be 0.9975. The model struggles to fit the QFM data using Bernstein and Trahiotis' parameters, because of the lower and steeper filter cut-off and the model has compensated for this by adopting a very low threshold value of  $\rho_i$ .

## 2. Effect of flanking bands

Flanking bands had very little mean effect on the QFM stimuli (though there were individual differences), but they increased thresholds substantially for the AM stimuli. The predictions of the peripheral-transduction model for the stimuli with flanking bands are shown in Fig. 4, without altering the parameters from the analysis above. The model still provides a good fit to the QFM data, but systematically underestimates thresholds in the AM case. The model cannot simultaneously account for data from stimuli both with and without flanking bands using the same parameters, because the peripheral non-linearities embodied in the model are insufficient to account for the effect of flanking bands. The model must be elaborated or replaced. Given the success of the peripheral transduction model in accounting for the data with no flanking bands, it is worth considering whether other known features of peripheral transduction could explain the present data. The most obvious candidate to account for an across-frequency non-linearity is suppression, in which the presence of an off-frequency tone reduces the response to a tone at the characteristic frequency of a given auditory nerve fiber ([Sachs and Kiang, 1968](#)). The flanking bands might thus be expected to suppress the response to the target band. However, it seems unlikely that this effect can account for impaired encoding of amplitude modulation, because

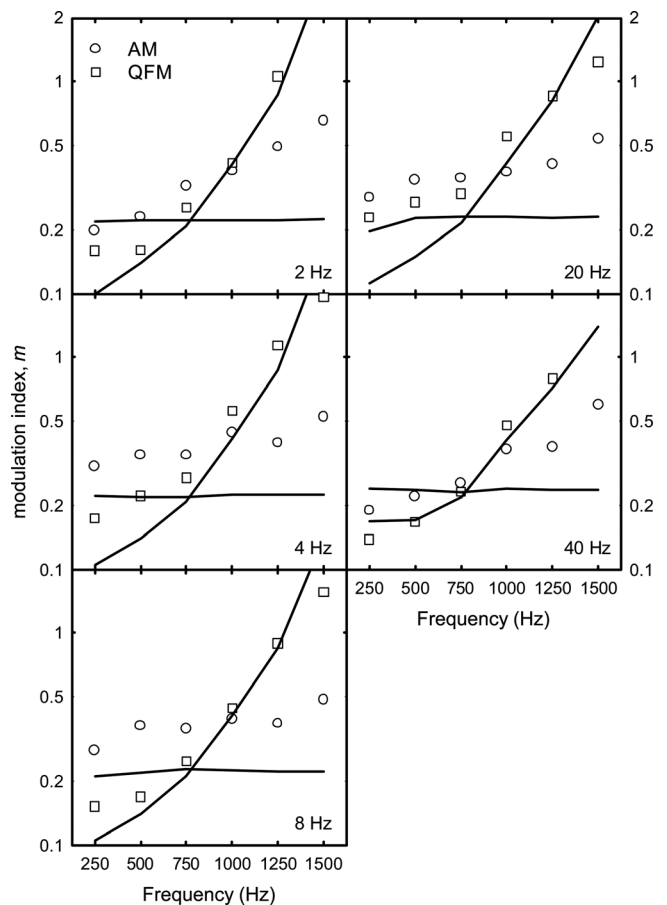


FIG. 4. Mean data from the conditions with flankers (symbols) plotted against the mean predictions of the same stimuli of a model (black lines) based on that of [Bernstein and Trahiotis \(2003\)](#), but using parameters fitted to the data for no flankers.

suppression is thought to be caused by a reduction in the gain of the active mechanism, as though the sound intensity has increased. The fact that AM thresholds were higher than QFM thresholds at low frequencies implies that the stimuli were presented within the compressive portion of the basilar membrane input–output function (Ruggero and Rich, 1991). The encoding of AM was thus already impaired and the addition of flanking bands would merely result in an upward shift in operating point within this compressive region.

The results thus seem inconsistent with the operation of a single mechanism for the detection of both ILD and ITD modulation. Instead, it appears that there may be one mechanism that is most sensitive to modulation of ILD and is susceptible to across-frequency interference and a second mechanism that is most sensitive to modulation of ITD, but is relatively unaffected by across-frequency interference. While flanking bands tended to increase interaural AM thresholds at high center frequencies, they had the opposite effect on interaural QFM thresholds. The reasons for this effect are unclear but its existence offers further evidence that there are two mechanisms with different characteristics.

The two-mechanism interpretation may also offer an account of other known phenomena. Individual differences in sensitivity to ILD modulation may explain those observed for high-frequency narrowband binaural unmasking, because high-frequency unmasking is dependent on the ILD sensitive mechanism. Binaural unmasking in narrowband maskers may be independent of frequency because it is dominated by the ILD sensitive mechanism, while that for broadband maskers is frequency-dependent because it is more dependent on the ITD sensitive mechanism. Finally, across-frequency interference has frequently been reported in the binaural literature, including the literature on binaural detection (Bernstein, 1991; Bernstein and Trahiotis, 1993, 1995), although the majority of reports concern lateralization tasks. Interestingly, the effect on binaural detection is asymmetric in frequency, with low frequency interferers disrupting detection of higher frequency targets, but not vice versa. Given that adding tones to noise in NOS $\pi$  gives rise to both fluctuations in interaural phase and amplitude, this asymmetry is also consistent with the data presented here. At low frequencies, the binaural system is most sensitive to interaural phase modulation, whose detection is unaffected by flanking bands; so at low frequencies the binaural signal detection is immune to across-frequency interference. At higher center frequencies (>750 Hz), the binaural system is more sensitive to interaural amplitude modulations, whose detection is affected by flanking bands; so at higher center frequencies the binaural signal detection is affected by across-frequency interference.

### 3. Effect of modulation rate

QFM thresholds were higher at high center frequencies when the modulation rate was low. At lower modulation rates, one would expect the binaural system to begin to temporally resolve the moving image of the target band. The fact that this improvement in temporal resolution of the lateral position coincided with elevated thresholds suggests

that lateral position was not the cue used by participants in the experiment. Alternatively, as modulation period begins to exceed the binaural temporal window (Culling and Summerfield, 1998), the interaural correlation within the window is less affected by the more limited amount of change in interaural phase that occurs within the window's span. Thus, differences in interaural correlation between the modulated and unmodulated intervals would be smaller. Although this result is consistent with the use of interaural correlation as a cue, it is not clear why the effect is observed only at high center frequencies. It is also possible that listeners may begin to exploit FM-induced AM as modulation rate increases, because auditory filters centered on the sidebands display such AM.

### 4. Underlying mechanisms

If two mechanisms are required to explain the data, then what would those mechanisms be? The data suggest only that one mechanism is most sensitive to ILD modulation, but is vulnerable to across-frequency interference, while the other is most sensitive to ITD modulation, but is invulnerable to such interference. However, the peripheral transduction model offers us a clear account of why sensitivity to ITD modulation might be greater at low frequencies regardless of the detection mechanism. Therefore, the dominant mechanism at low frequencies may well be one that is simply sensitive to changes in interaural correlation (explicitly so in the case of correlation-based models, or implicitly so in the case of E-C theory). On the other hand, the mechanism more sensitive to ILD modulation may be specifically tuned to this cue, but integrates information across frequency, resulting in its susceptibility to interference.

## III. EFFECT OF PERCEPTUAL GROUPING

One conclusion from the main experiment was that a two-mechanism model might offer an explanation of binaural interference in binaural unmasking. Binaural interference has most commonly been studied with respect to detection of static differences in interaural intensity or time delay, where it has been found to depend on grouping cues. For instance, Best *et al.* (2007) found that embedding the interfering sound within a sequence of similar sounds resulted in “capture” of the interferer into a separate sound stream and removed its influence from the detection of an interaural time delay applied to a target sound. A few studies have also examined binaural interference in binaural unmasking tasks, but here the effects of grouping cues appear more equivocal.

Bernstein (1991) measured BMLD for an 800-Hz signal masked in the NOS $\pi$  binaural configuration by continuous broadband noise. An interfering 400-Hz tone that was gated with the 800-Hz signal tone produced substantial interference, but a similar continuous tone did not, suggesting that the interference effect was dependent upon grouping by common onset and offset time with the target tone. On the other hand, Bernstein and Trahiotis (1993) found that when the interfering tone was also pulsed, but preceded the onset of the target tone by up to 320 ms, this onset asynchrony which would normally be expected to provide powerful

segregation (e.g., Darwin, 1981) did not reduce the interference effect. Finally, Bernstein and Trahiotis (1995) used a 4-kHz signal masked by a 100- or 400-Hz wide band of noise in the context of a 100- or 400-Hz wide band of interfering noise, centered at 500 Hz; here, the interference effect still depended upon whether the interferer was continuous or gated with the other components of the stimulus. Thus some results suggest that the interference effect is removed if the interferer is not temporally coterminous with the signal, while others do not.

A follow-up experiment was therefore conducted in order to assess whether the effect of flanking bands on detection of interaural AM could also be attenuated by introducing differences in the gating of target and interferer.

### A. Stimuli

The stimuli were similar to those of the main experiment aside from manipulations of the flanking bands. Four conditions were prepared: In condition “none” the flankers were absent; in condition “gated” the flankers were gated coterminously with the target band; in condition “fringed” the flanking band onset was 300 ms before that of the target band, but the two offset simultaneously; in condition “continuous” the flanking bands were continuous. The first three conditions were prepared using similar methods to the main experiment, entirely within MATLAB. In the fringed condition, 10-ms, raised-cosine onset ramps were applied separately to the target and flanking bands. The continuous flanking bands of the continuous condition were created in Cool Edit '96, a program for digital generation and editing of sounds. A 30-s buffer of Gaussian noise (44.1-kHz sampling rate and 16-bit sample depth) was filtered with the FFT-filter function to the same specifications as the MATLAB-generated flanking bands. The buffer was then spliced into a continuous loop using a cross-fading technique. No audible artifact was heard during the 100-ms cross-fade. The sound level was then adjusted using a sound level meter to match within 1 dB that produced by the MATLAB-generated flanking bands. Thresholds were collected in the continuous condition by playing the resulting loop continuously while running the MATLAB program as for condition “none.” Sounds from MATLAB and CoolEdit were thus mixed digitally by the Windows XP operating system<sup>1</sup>.

### B. Procedure

Three listeners took part in the experiment. Listeners JC and NC were the same as those who took part in the main experiment. One new listener, RB, took to the task very rapidly and was able to produce lower thresholds than the other two listeners after only two adaptive tracks of practice. Eight conditions were tested, all with a target-band center frequency of 500 Hz. These included two modulation types (QFM/AM) and four flanker types (none, gated, fringed, and continuous). Since an experimental session was broken up by the need to manually start and stop the continuous flanking bands, the different conditions were manually controlled in an *ad hoc*, pseudorandom order. Three thresholds were

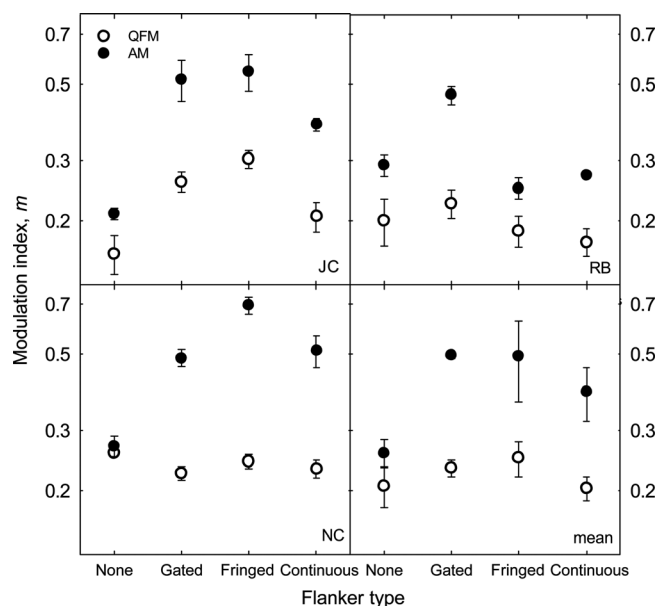


FIG. 5. Mean data from experiment 2. Error bars are one standard error of the mean.

collected for each of the eight conditions over the course of three 1-h sessions.

### C. Results

Figure 5 shows the mean thresholds for each individual listener (first three panels) and their mean (bottom right panel). The variance of individual mean thresholds was small. Despite the small sample of listeners, a wide spectrum of individual differences is apparent. Consistent with the main experiment, all listeners showed lower thresholds for QFM than for AM for a target band centered at 500 Hz. All listeners also showed some effect of interference from the flankers in the AM/gated condition. Other effects were more diverse, so each listener’s data was analyzed separately using analysis of variance and Tukey pairwise comparisons.

Listener JC (top left panel) had higher thresholds when flankers were present (interference) for both AM and QFM, where those flankers were gated with the target band or had a 300-ms leading temporal fringe [ $F(1,4) = 23, p < 0.001; q = 9.3, p < 0.01; q = 10.9, p < 0.01$ ]. The effect was not significantly larger for AM. It should be noted, however, that in the main experiment JC had shown no evidence of interference in the QFM case. A 300-ms temporal fringe does not reduce the interference, but thresholds were somewhat lower (for both modulation types) when the flanking bands were continuous ( $q = 5.9, p < 0.05$ ). Even in the continuous case, thresholds did not return to the levels seen in the absence of flanking bands.

Listener NC (bottom left panel) only showed interference for AM, resulting in main effects of both modulation type and flanker type [ $F(1,2) = 61, p < 0.02; F(3,6) = 31, p < 0.001$ ], but also an interaction between the two [ $F(3,6) = 34, p < 0.001$ ]. Tukey comparisons confirmed that there were no significant differences between mean



thresholds for different flanker types in the QFM case. In the AM case, thresholds were higher for each of the three cases in which flankers were present, than when they were absent ( $q < 9$ ,  $p < 0.01$  in each case). Interference was greatest in the fringed case; this condition differs significantly from the gated and continuous cases ( $q > 8$ ,  $p < 0.05$ , in each case). Interference was thus reduced with continuous flanking bands, but only compared to the fringed case.

Like NC, RB (top right panel) only showed interference for AM, resulting in significant main effects [ $F(1,2) = 44$ ,  $p < 0.05$ ;  $F(1,2) = 31$ ,  $p < 0.001$ ] and interaction [ $F(3,6) = 5.1$ ,  $p < 0.05$ ]. In this case, however, interference was completely eliminated either by the temporal fringe or by playing the flanking bands continuously. Consequently, there were significant differences only between the AM/gated case and all other conditions ( $q > 7$ ,  $p < 0.05$  in each case).

Looking at the mean thresholds across listeners (bottom right panel), there appears little evidence of interference for the QFM modulation type. The diversity of effects generated by the fringed and continuous flankers for the AM modulation type is clearly shown by the large error bars for these two conditions.

#### D. Discussion

The purpose of the experiment was to determine whether the interference effect found in the main experiment is the same as that described in the literature, and therefore, that it exhibits the same characteristics. Given the results reported by [Bernstein \(1991\)](#) and [Bernstein and Trahiotis \(1993\)](#), one might have expected to see that adding a 300-ms forward temporal fringe to the flanking bands would have had no effect, while playing the flankers continuously would have eliminated the interference effect. This prediction is broadly consistent with the results of JC and NC, in that both these listeners showed a partial abolition of the effect only when continuous flankers were presented. However, listener RB confounded such expectations, as her results showed a complete abolition of the effect for both the fringed and the continuous flanker conditions. The clear evidence of diversity observed here, coupled with the small numbers of participants in all of the relevant studies (three in [Bernstein, 1991](#), four in [Bernstein and Trahiotis, 1993](#) and three here) leaves open the possibility that there is in fact no characteristic signature by which one can recognize a common interference mechanism across different experiments. In order to verify a common pattern it will be necessary to use the same participants across all of the different paradigms in the same experiment.

[Bernstein and Trahiotis \(2004\)](#) reported that the lateralization of high-frequency transposed stimuli appeared to be immune to binaural interference. Although the transposed stimuli featured interaural differences in the amplitude envelope, these differences were designed to simulate the encoding of fine-structure differences occurring at lower frequencies. Consequently, they probably have more in common with the QFM stimuli used in the present experiment than the AM stimuli. The lack of interference in their experiment may thus be consistent with the lack of interference

generally observed in the QFM case, here, and be a feature of temporally encoded interaural differences.

#### IV. CONCLUSIONS

Sensitivity to modulation of ITD and ILD in a narrow-band noise showed clear patterns of difference. These differences occurred on two levels. First, sensitivity to modulated ITDs (QFM) deteriorated with increasing center frequency, while sensitivity to modulated ILDs (AM) was largely unaffected. Second, sensitivity to modulated ILDs was sensitive to interference from unmodulated noise elsewhere in the frequency spectrum, while sensitivity to modulated ITDs was largely unaffected. A model of peripheral transduction that compresses the stimulus envelope and progressively removes fine structure information with increasing frequency can account for the first effect. The effect is thus consistent with the idea that both forms of interaural modulation are detected by a mechanism sensitive to the internal interaural correlation (the correlations between the stimulus waveforms after nonlinear peripheral transduction processes). However, the fact that the second effect only occurred for ILD modulation cannot be explained by such a model, suggesting that different detection mechanisms must be at work. All listeners are able to overcome this interference to some extent using perceptual grouping, suggesting that encoding of modulated ILDs and ITDs is separate even at quite high levels of perceptual processing. This ability was highly variable across listeners, perhaps explaining apparent inconsistencies between previous reports.

#### ACKNOWLEDGMENTS

The author is indebted to the editor, Ruth Litovsky and two anonymous reviewers for their helpful suggestions in refining this manuscript.

#### APPENDIX

The MATLAB iterative multi-parameter search function `fminsearch` was first used to optimize the parameters of the [Bernstein and Trahiotis \(2003\)](#) transduction model (filter cut-off, compression exponent and  $\rho_i$  threshold) for the current data set. Filter orders between three and five were tested exhaustively in separate parameter searches. A second, nested `fminsearch` was used to find the predicted threshold value of  $m$  for each condition. This search adjusted  $m$  such that it minimized the difference between  $\rho_i$  and the current threshold value across all 120 experimental conditions [5 modulation rates  $\times$  6 center frequencies  $\times$  2 modulation types  $\times$  2 modes (with flankers/no flankers)]. In order to avoid stochastic variation, all stimuli in a given search were based upon a single 2-s noise sample which was used repeatedly to create 2-s stimuli with adjusted experimental parameters. Repeated runs of the search showed that a 2-s stimulus duration was sufficient to generate consistent results to 2–3 significant digits. In other respects, stimuli were synthesized exactly as in the experiment, before being processed by the transduction model. The transduction model consisted of (1) a gammatone filter tuned to the center-frequency of the target band, (2) Hilbert transformation, (3) compression of the

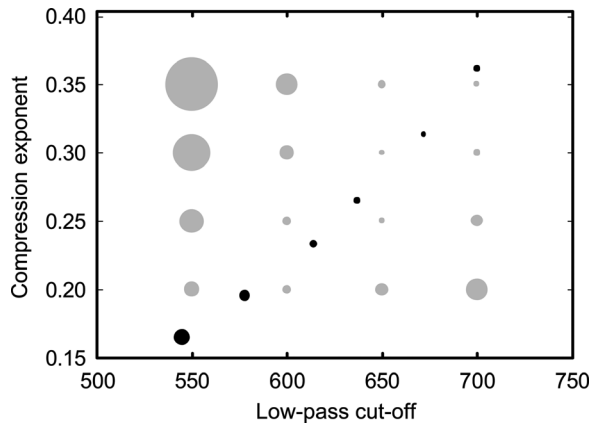


FIG. 6. Log mean-squared error of the peripheral-nonlinearity model fit is represented by symbol size as a function of low-pass cut-off frequency and compression exponent. Gray symbols are fits with fixed values of cut-off frequency and compression exponent in which only threshold internal interaural correlation was fitted. Black symbols are different fits in which all three parameters were fitted. The parameters from the third of these points were selected for use in generating Fig. 3 due to the close correspondence of the fitted compression factor with that used by Bernstein and Trahiotis (2003).

resulting envelope using the compression exponent, (4) inverse Hilbert transformation, (5) half-wave rectification, and (6) low-pass filtering with a Butterworth filter. Following Bernstein and Trahiotis (1996),  $\rho_i$  was evaluated using the normalized correlation between the transduced signals for each ear. The fitted parameters at the end of the fitting process were recorded.

At an early stage, it was found that the third-order filters tended to give closer fits than higher orders, but the other parameters proved difficult to optimize. This difficulty was caused by an awkward search space which featured a deep valley with local minima along its floor. This valley is illustrated in Fig. 6. The diameters of the gray circles in Fig. 6 represent the log sum-squared error for different fits in which the filter cut-off and the compression exponent were fixed and only the threshold  $\rho_i$  was adjusted: Smaller circles are thus closer to the valley floor. It can be seen that the optimal cut-off varies as a function of the compression exponent. The optimal  $\rho_i$  threshold also changes as one moves along this valley from 0.993 at 550 Hz cut-off to around 0.98 at 700 Hz cut-off. Black circles similarly represent the log sum-squared error for fits in which all three parameters were freely varied. These points give the best estimate of the exact location of the valley floor.

<sup>1</sup>Separate evaluation of the effect of mixing signals from different applications in this way indicated that distortion was not introduced when, as in this case, all applications employed a common sampling rate and the combined signal does not exceed the 16-bit bit-depth of the digital-to-analog converter.

Bernstein, L. R. (1991). "Spectral interference in a binaural detection task," *J. Acoust. Soc. Am.* **89**, 1306–1313.  
 Bernstein, L. R., and Trahiotis, C. (1993). "Spectral interference in a binaural detection task: Effects of masker bandwidth and temporal fringe," *J. Acoust. Soc. Am.* **94**, 735–742.  
 Bernstein, L. R. and Trahiotis, C., (1995). "Binaural interference effects measured with masking-level difference and with ITD- and IID-discrimination paradigms," *J. Acoust. Soc. Am.* **98**, 155–163.

Bernstein, L. R., and Trahiotis, C. (1996). "The normalized correlation: Accounting for binaural detection across center frequency," *J. Acoust. Soc. Am.* **100**, 3774–3784.  
 Bernstein, L. R., and Trahiotis, C. (2003). "Enhancing interaural-delay-based extents of laterality at high frequencies by using 'transposed stimuli'," *J. Acoust. Soc. Am.* **113**, 3335–3347.  
 Bernstein, L. R., and Trahiotis, C. (2004). "The apparent immunity of high-frequency 'transposed' stimuli to low-frequency binaural interference," *J. Acoust. Soc. Am.* **116**, 3062–3069.  
 Bernstein, L. R., Trahiotis, C., and Hyde, E. L. (1998). "Inter-individual differences in binaural detection of low-frequency or high-frequency tonal signals masked by narrow-band or broadband noise," *J. Acoust. Soc. Am.* **103**, 2069–2078.  
 Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (2007). "Binaural interference and auditory grouping," *J. Acoust. Soc. Am.* **121**, 1070–1076.  
 Colburn, H. S., and Durlach, N. I. (1978). Models of binaural interaction, in *Handbook of Perception*, edited by E. C. Carterette, (Academic, New York, 1978), Vol. IV.  
 Colburn, H. S., (1977). "Theory of binaural interaction based on auditory nerve data II: Detection of tones in noise," *J. Acoust. Soc. Am.* **61**, 525–533.  
 Colburn, H. S., (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.  
 Culling, J. F. (2007). "Evidence specifically favoring the equalization-cancellation theory of binaural unmasking," *J. Acoust. Soc. Am.* **122**, 2803–2813.  
 Culling, J. F., and Summerfield, Q. (1995). "Perceptual separation of concurrent speech sounds: Absence of across-frequency grouping by common interaural delay," *J. Acoust. Soc. Am.* **98**, 785–797.  
 Culling, J. F., and Summerfield, Q. (1998). "Measurements of the binaural temporal window using a detection task," *J. Acoust. Soc. Am.* **103**, 3540–3553.  
 Darwin, C. J. (1981). "Perceptual grouping of speech components differing in fundamental frequency and onset time," *Q. J. Exp. Psychol. A* **33**, 185–207.  
 Davidson, S. A., Gilkey, R. H., Colburn, H. S., and Carney, L. H. (2009). "Diotic and dichotic detection with reproducible chimeric stimuli," *J. Acoust. Soc. Am.* **126**, 1889–1905.  
 Durlach, N. I. (1963). "Equalization and cancellation theory of binaural masking level differences," *J. Acoust. Soc. Am.* **35**, 1206–1218.  
 Durlach, N. I. (1972). "Binaural signal detection: equalization and cancellation theory" in Tobias, J. V. (Ed.) *Foundations of Modern Auditory Theory*. Vol II (Academic, New York) pp. 369–462.  
 Durlach, N. I., Gabriel, K. J., Colburn, H. S., and Trahiotis, C. (1986). "Interaural correlation discrimination: II. Relation to binaural unmasking," *J. Acoust. Soc. Am.* **79**, 1548–1557.  
 Goupell, M. J., and Hartmann, W. M. (2006). "Interaural fluctuations and the detection of interaural incoherence: Bandwidth effects," *J. Acoust. Soc. Am.* **119**, 3971–3986.  
 Goupell, M. J., and Hartmann, W. M. (2007). "Interaural fluctuations and the detection of interaural incoherence. III. Narrowband experiments and binaural models," *J. Acoust. Soc. Am.* **122**, 1029–1045.  
 Hafter, E. R., and Carrier, S. C. (1970). "Masking-level differences obtained with a pulsed tonal masker," *J. Acoust. Soc. Am.* **47**, 1041–1047.  
 Hirsh, I. J. (1948). "The influence of interaural phase on interaural summation and inhibition," *J. Acoust. Soc. Am.* **20**, 536–544.  
 Hirsh, I. J., and Burgeat, M. (1958). "Binaural effects in remote masking," *J. Acoust. Soc. Am.* **30**, 827–832.  
 Koehnke, J., Colburn, H. S., and Durlach, N. I. (1986). "Performance in several binaural-interaction experiments," *J. Acoust. Soc. Am.* **79**, 1558–1562.  
 Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.  
 McFadden, D., and Pasanen, E. G. (1974). "High-frequency masking-level differences with narrow-band noise signals," *J. Acoust. Soc. Am.* **56**, 1226–1230.  
 McFadden, D., Jeffress, L. A., and Ermev, H. L. (1971). "Differences of interaural phase and level in detection and lateralization: 250 Hz," *J. Acoust. Soc. Am.* **50**, 1484–1493.  
 Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.  
 Osman, E. (1971). "A correlation model of binaural masking level differences," *J. Acoust. Soc. Am.* **50**, 1494–1495.  
 Patterson, R. D., Allerhand, M. H., and Gigu re, C., (1995). "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Am.* **98**, 1890–1894.

- Ruggero, M. A., and Rich, N. C. (1991). "Furosemide alters organ of corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane," *J. Neurosci.* **11**, 1057–1067.
- Sachs, M. B., and Kiang, N. Y.-S. (1968). "Two-tone inhibition in auditory nerve fibres," *J. Acoust. Soc. Am.* **43**, 1100–1128.
- Stern, R. M., and Colburn, H. S. (1985). "Lateral-position-based models of interaural discrimination," *J. Acoust. Soc. Am.* **77**, 753–755.
- van de Par, S., and Kohlrausch, A. (1998). "Diotic and dichotic detection using multiplied-noise maskers," *J. Acoust. Soc. Am.* **103**, 2100–2110.
- van der Heijden, M., and Joris, P. X. (2010). "Interaural correlation fails to account for detection in a classic binaural task: Dynamic ITDs dominate N0S $\pi$  detection," *J. Assoc. Res. Otolaryngol.* **11**, 113–131.
- Yost, W. A. (1972). "Tone-on-tone masking for three binaural listening conditions," *J. Acoust. Soc. Am.* **52**, 1234–1237.