No spatial release from amplitude modulation masking

Eric R. Thompson[1,2], Torsten Dau[1] and Barbara G. Shinn-Cunningham[2]

1 *Centre for Applied Hearing Research, Technical University of Denmark, DK-2800 Lyngby, Denmark*

2 *Hearing Research Center, Boston University, Boston, MA 02215, USA*

In many signal detection and speech intelligibility studies, performance is improved if there is a perceived spatial separation between the target and the masker, as compared to when the target and masker are perceived to be co-located. The goal of the present study was to determine if a similar spatial release from masking can be measured in a masked amplitude modulation (AM) detection experiment. Temporally interleaved transposed stimuli were used as AM carriers for the experiments. These carriers could be lateralized separately using only interaural time differences (ITDs) and not interaural level differences (ILDs). The first experiment measured the perceived lateral positions of the probe and masker, independently and in combination. Experiment two used the same stimuli as carriers for measuring masked modulation detection thresholds. The results showed that the probe and masker could be perceived to come from separate lateral positions. The modulation detection results showed bandpass modulation-frequency tuning. There was no improvement in modulation thresholds when the target was diotic and the masker was lateralized with a 1-ms ITD, as compared to the co-located condition.

INTRODUCTION

In many masked signal detection experiments, there can be a significant improvement (often around 15 to 20 dB) in the detection thresholds when there is a perceived spatial separation between the target and the masker, created using interaural time differences (ITD). This has been demonstrated in the masked detection of signals (e.g., van der Heijden and Trahiotis, 1999; Kopčo and Shinn-Cunningham 2008). These prior studies all dealt with the ability to hear a sound presented simultaneously with a masking noise, but were not concerned with the ability to hear details about the target sound itself, e.g., its pitch trajectory or the shape of its amplitude envelope. The present study sought to determine whether there is a similar spatial release from masking in the detection of sinusoidal amplitude modulation (AM) imposed on a suprathreshold carrier in the presence of an amplitude-modulated masker.

Previous studies of masked AM detection have focused on monaural or diotic listening (e.g., Houtgast, 1989). In those prior studies, a sinusoidal AM target was imposed together with an AM masker (either sinusoidal or narrowband noise) on a broadband-noise or pure-tone carrier, and the minimum target modulation depth was measured as a function of the modulation-frequency spectrum of the masker. These experiments typically show modulation-frequency tuning in that the thresholds are

Proceedings of ISAAR 2009: Binaural Processing and Spatial Hearing. 2nd International Symposium on Auditory and Audiological Research. August 2009, Elsinore, Denmark. Edited by J. M. Buchholz, T. Dau, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN: 87-990013-2-2. The Danavox Jubilee Foundation, 2010.

highest when the modulation frequency of an AM target is close to or within the modulation-frequency spectrum of the AM masker, and lower when the target and masker are separated in their modulation spectra.

Other studies have investigated AM detection and binaural interactions in terms of binaural modulation detection interference (MDI; Bacon and Opie, 1994; Sheft and Yost, 1997). MDI is a form of masked modulation detection where the target AM and masker AM are imposed on carriers in different audio-frequency regions. For example, when the target AM was applied to a 1-kHz carrier and a masker AM was applied to a 4-kHz carrier, and both were presented monaurally in the same ear, then thresholds increased by about 7 to 8 dB relative to the unmasked case. When the masker was moved to the other ear, then thresholds were only about 2 to 3 dB higher than in the unmasked case (Bacon and Opie, 1994), showing about a 5 dB release from masking. When the target and masker modulation were presented with different ITDs, so that there was target and masker energy in both ears (Sheft and Yost, 1997), there was an improvement of about 2 dB in the thresholds relative to the diotic condition, showing a small, but significant, release from MDI. These studies suggest that there is a small interaction between AM processing and binaural (spatial) processing in the auditory system, at least across audio-frequency channels.

The goal of the experiments in the present study was to investigate the interactions between AM and binaural processing by measuring lateralization of AM carriers and masked AM detection thresholds when the probe and masker AM were imposed on carriers within the same spectral region (similar to the experiments from Houtgast, 1989), but whose ITDs could be controlled independently.

GENERAL METHODS

The two experiments performed in the present study measured different aspects of perception using the same stimuli. The first experiment tested whether two AM carriers with the same spectral content could be lateralized separately using ITDs only. When the two carriers could be lateralized separately, the second experiment measured AM detection thresholds as a function of the masker ITD and modulation-frequency content. The common aspects of the experiments are presented here, with specific details in the subsequent sections.

Stimuli

The main requirement for the stimuli was to have two temporally and spectrally overlapping carriers be perceived to come from different lateral positions. A sound can be lateralized by creating an ILD and/or an ITD. By definition, an ILD changes a signal's energy in at least one ear, thereby creating different SNRs in each ear when the target and masker have different ILDs. In many cases, signal detection thresholds with ILD-based lateralization can simply be predicted by considering the ear with the higher SNR, or the "better-ear advantage." Therefore, for the present study, it was desired to only use ITDs to create the lateralization.

In order to have different ITDs on the probe and masker carriers without phase cancellation effects, temporally-interleaved transposed stimuli were used as the carriers. Transposed stimuli were originally designed to create firing patterns in high-frequency auditory-nerve fibers, based on the stimulus envelope, that are similar to those of low-frequency nerves, based on the stimulus fine-structure (van de Par and Kohlrausch, 1997).

All stimuli were generated digitally using MATLAB at a sampling rate of approx. 98 kHz. The signal envelope was created by convolving the positive half-wave of a 250-Hz sinusoid with a 125-Hz pulse train. In order to restrict the bandwidth of the signal, it was low-pass filtered at 1250 Hz. This envelope was then multiplied with a 5-kHz sinusoidal carrier, creating the transposed stimulus. A second stimulus was created using the same method, only with a time delay that was randomly drawn from a distribution given by 4±B(4,2) [ms], where B(4,2) is a beta distribution, with equal probability of plus or minus. This distribution was selected so that the two stimuli's pulses would have no temporal overlap for any combination of ITDs up to ±1 ms. An ITD and AM were then applied to each stimulus, depending on the measurement condition. An excerpt of the result of adding the two stimuli together is shown in Fig. 1, with the probe stimulus in black and the masker stimulus in gray.

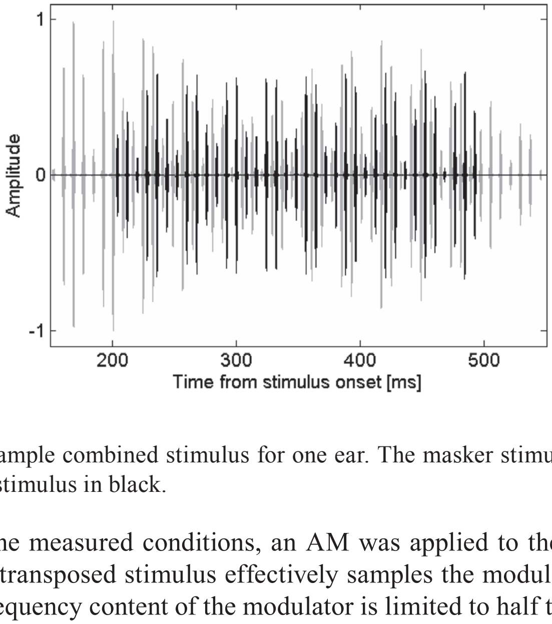

Fig. 1: Example combined stimulus for one ear. The masker stimulus is in gray and the probe stimulus in black.

In some of the measured conditions, an AM was applied to the masker and probe stimuli. The transposed stimulus effectively samples the modulator at a rate of 125 Hz, so the frequency content of the modulator is limited to half this rate, or 62.5 Hz, to avoid aliasing. The probe stimulus was defined as a function of time t as:

$$y_{probe}(t) = [1 + m\sin(2\pi f_m t + \varphi)] \cdot x_{probe}(t) \tag{Eq. 1}$$

where m is the modulation depth, f_m is the modulation frequency (always 16 Hz in the present study), φ was a random starting phase selected from a uniform distribution over the interval $[0; 2\pi]$, and $x_{probe}(t)$ is the transposed-stimulus carrier described above. The masker modulator was a 5.6-Hz-wide Gaussian-noise masker with a variance of 0.1 (power of 10 dB FS) and was applied as:

$$y_{mask}(t) = [1 + n(t)] \cdot x_{mask}(t) \qquad \text{(Eq. 2)}$$

where $n(t)$ is the masker modulator and $x_{mask}(t)$ is the transposed-stimulus carrier described above. The center frequency of the masker noise was an experimental parameter, as described below. The masker modulator was generated by creating a two-second-long Gaussian noise and setting the magnitudes of the frequency components outside the desired passband to zero.

In order to further enhance the perceptual separability of the two stimuli, they were also gated separately. Both stimuli were gated with 10 ms $\cos^2$ ramps at the beginning and end. The masker stimuli had an overall duration of 600 ms, and the probe stimuli had a 200 ms gating delay, relative to the onset of the masker stimulus, and an overall duration of 300 ms. An excerpt from an example of a combined stimulus is shown in Fig. 2.

Next, an ITD was applied to the masker and probe stimuli in the time domain. Then, the levels of the two stimuli were equalized, they were added together to form the complete stimulus, and the overall level was set to 65 dB SPL in each ear.

EXPERIMENT I: ITD-BASED LATERALIZATION OF MASKER AND PROBE

Procedure

The lateralization experiment was conducted to determine whether the listeners would be able to lateralize the masker and the probe stimuli separately, or whether the two temporally-interleaved stimuli could be heard at two separate locations. The listeners' task was to adjust the ILD of a pointer sound until it was perceived to come from the same lateral position within the head as the target stimulus, which was either the masker or the probe stimulus, depending on the condition. The ILD-pointer method was similar to methods used in previous studies to measure the extent of laterality of a stimulus (e.g., Bernstein and Trahiotis, 2003). There were six conditions measured: (1) the alignment target was the probe stimulus, presented alone, (2) the target was the masker stimulus, presented alone, (3) the target was the probe stimulus with a fixed ITD of 0 ms, presented in the combined stimulus, measured as a function of the masker ITD, (4) the target was the masker stimulus, presented in the combined stimulus, measured as a function of the masker ITD when the probe had a fixed ITD of 0 ms, (5) the target was the probe stimulus, presented in the combined stimulus, measured as a function of the probe ITD when the masker had a fixed ITD of 1 ms, and (6) the target was the masker with a fixed ITD of 1 ms, presented in the combined stimulus, measured as a function of the probe ITD (the conditions are also summarized in the legend of Fig. 2). Note that the stimuli were identical in conditions 3 and 4, and in conditions 5 and 6, but the target for alignment changed between the conditions. These six conditions were measured as a function of either the masker or probe ITD τ for τ = {-1; -0.5; 0; 0:5; 1} ms, with a positive ITD indicating right ear leading. For this experiment, the masker AM was always centered at 16 Hz, the probe AM frequency.

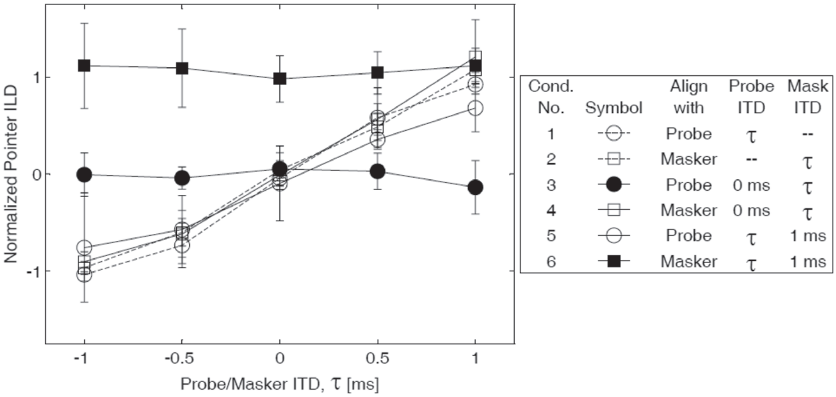

Fig. 2: Normalized ILD pointers as a function of masker or probe ITD. The circles indicate alignment with the probe stimulus, and the squares indicate alignment with the masker stimulus. The data points connected with dashed lines were measured with one stimulus only, and those connected with solid lines were measured as part of the combined stimulus. The solid black markers indicate that the points were measured for a fixed ITD (see legend) as a function of the ITD (τ) of the other part of the combined stimulus.

The ILD pointer was a 5-kHz tone, fully amplitude modulated at 250 Hz. The duration of the pointer was 600 ms, and was gated with 10 ms $\cos^2$ ramps. Each measurement started with a random pointer ILD selected from a uniform distribution of integers from -8 to +8 dB, with a positive ILD indicating a higher level in the right ear. The probe, masker or combined stimulus was presented first, followed by a 200-ms pause and the pointer stimulus. The listener could step the pointer to the right or left by changing its ILD or continue to the next alignment target when they were satisfied with the alignment. The stimuli were presented in blocks by the condition, with conditions 3 and 5 blocked together and conditions 4 and 6 blocked together, with the five ITDs for each condition presented in random sequence within each block. Each test subject completed five repetitions for each data point.

Results

Figure 2 shows the normalized pointer ILD required to align with the lateral position of the stimuli in the six measured conditions as a function of the ITD parameter τ. There was a large variance between test subjects in the pointer ILD required for alignment, similar to prior studies (e. g., Heller and Trahiotis, 1996), and may reflect a different relative sensitivity to ILDs and ITDs between subjects. Therefore, each test subject's data were normalized for plotting and for the analysis using their mean data from the lateralizations when the two stimuli were presented alone (conditions 1 and 2). The results shown in Fig. 2 are the mean normalized data across subjects with error bars representing one standard deviation of the mean. The data were analyzed in the following using analyses of variance with repeated measures (RM-ANOVA) with a level of 0.05 required for statistical significance.

The two conditions with a fixed ITD on the alignment target (conditions 3 and 6, filled symbols in Fig. 2), where the probe lateralization was measured as a function of the masker ITD and vice versa, show little change with the ITD of the other stimulus. In condition 3 (filled circles in Fig. 2), the ITD of the probe stimulus was fixed at 0 ms, and there was no significant change in the normalized pointer ILD as the ITD of the masker stimulus was changed from -1 to 1 ms [$F(4,16)=0.56$, $p=0.70$]. In condition 6 (filled squares in Fig. 2), the ITD of the masker stimulus was fixed at 1 ms, and there was also no significant change in the normalized pointer ILD with the ITD of the probe stimulus [$F(4,16)=0.88$, $p=0.50$].

The other four conditions, for which the alignment target's ITD was the experimental parameter, all show a monotonic increase in normalized pointer ILD as the ITDs of the target stimuli were increased from -1 to 1 ms. An RM-ANOVA across these four conditions with main factors of condition and ITD found a significant effect of ITD [$F(4,16)=159$, $p<0.0001$], but no significant effect of condition [$F(3,12)=0.88$, $p=0.48$]. There was, however, a significant interaction [$F(12,48)=2.15$, $p<0.05$], which a post hoc analysis showed to be a small repulsion effect in condition 5 (open circles connected with a solid line in Fig. 2) for positive ITDs (see also, e.g., Carlile *et al.*, 2001). This repulsion was only seen in the data of two listeners.

Discussion

The results of the lateralization experiment show that the two stimuli (masker and probe) can be perceived as coming from two separate locations within the head. This suggests that they may be perceived as two separate auditory objects.

The two stimuli were gated separately, so the masker stimulus began 200 ms before the onset of the probe stimulus. Previous studies have shown robust lateralization with 100 ms and 200 ms long stimuli (Buell *et al.*, 1991), so it is likely that the listeners could lateralize the masker stimulus before the probe stimulus even started. On the other hand, the probe stimulus started after the masker stimulus and ended before the masker stimulus ended, so it could only have been lateralized separately from the masker stimulus if it was perceived to be a separate auditory object from the masker stimulus.

When two sounds are presented simultaneously with similar interaural parameters, but are not perceptually grouped together, there can be a repulsion effect, or pushing. This effect is seen in the data shown in Fig. 2 for condition 5 (open circles connected with a solid line) with positive ITDs, where some of the listeners required smaller pointer ILDs to align with the probe stimulus than were required for the same ITD in the other conditions (see conditions 1, 2 and 4 in Fig. 2), indicating that the perceived position was pushed towards the mid-line. Since there was no significant pushing effect seen with a masker ITD of 1 ms and a probe ITD of 0 ms (diotic), it was assumed that the perceived distance between the masker and probe was larger enough that they could be used to investigate a spatial release from modulation masking in a modulation-detection experiment.

EXPERIMENT II: MASKED AMPLITUDE MODULATION DETECTION

Procedure

The same temporally-interleaved transposed stimuli were used as carriers in the modulation detection experiment to test the hypothesis that a perceived spatial separation of a probe and a masker would improve the detectability of AM applied to the probe. Similar methods to those used in previous masked modulation detection studies were used (e.g., Houtgast, 1989), in which a narrowband-noise modulator was used to interfere with detection of a sinusoidal modulator. In the previous studies, the masker and probe modulators were applied in series to a common carrier. In the present study, in order to achieve a spatial separation of the masker and the probe, the modulators were applied to separate carriers.

The minimum modulation depth m required to detect a sinusoidal AM imposed on the probe stimulus was measured as a function of the center frequency of the masker modulator, and of the masker ITD, either 0 ms (diotic) or 1 ms. For this experiment, the probe stimulus always had a 0 ms ITD. The center frequencies of the masker were 6.3, 10.1, 16, 25.4 and 40.3 Hz. In addition, two control measurements were made with no AM imposed on the masker transposed stimulus, both with 0 ms ITD and with 1 ms masker ITD.

A 3-interval, 3-alternative, forced-choice (3-AFC) design was used, with one randomly selected signal interval and two reference intervals. There was no AM imposed on the probe stimulus in the reference intervals, and a sinusoidal AM was imposed on the probe stimulus in the signal interval. The modulation depth was adaptively tracked, following a 2-down, 1-up rule (Levitt, 1971). Each track started with a modulation depth of -5 dB, and a step size of 4 dB. After the second and fourth change of step direction, the step size was halved, and the track continued for eight further reversals at the final step size of 1 dB. The threshold was defined as the mean of the modulation depths (in dB) of the last eight reversals in each track. Each test subject completed four tracks for each data point.

Results

Figure 3a shows the modulation detection thresholds for four masker conditions: (1) with no AM imposed on the masker stimulus with the masker carrier at 0-ms ITD (filled circles), (2) with no AM imposed on the masker stimulus with the masker carrier at 1-ms ITD (filled squares), (3) with a noise AM imposed on the masker stimulus with a masker ITD of 0 ms (open circles), and (4) with a noise AM imposed on the masker stimulus with a masker ITD of 1 ms (open squares). Figure 3b shows the same data, but replotted as the difference in threshold between the conditions with the modulated and unmodulated maskers for corresponding masker ITDs.

The RM-ANOVA showed that a significant effect of masker ITD [$F(1,4)=10.5$, $p<0.05$], and of masker center frequency [$F(4,16)=10.3$, $p<0.001$], but no significant interaction [$F(4,16)=2.04$, $p=0.14$]. Post-hoc analyses showed that the thresholds measured with a 0-ms masker ITD were lower than those with a 1-ms masker ITD. The data from the unmodulated-masker conditions showed no significant effect of masker ITD on the thresholds [$F(1,4)=1.54$, $p=0.28$].

Discussion

The data do not show any spatial release from modulation masking. Even though the masker and probe could be lateralized separately, indicating a perceptual segregation, the modulation detection thresholds did not improve. In fact, there was a small increase in the modulation detection thresholds when the masker was at a different perceived lateral position than the probe. However, this did not translate into a statistically significant increase in the modulation masking levels. These results were unexpected from the original hypothesis that there would be a spatial release from the modulation masking, in which case, the detection thresholds with a masker ITD of 1 ms would have been lower (less masked) than the detection thresholds measured with a diotic masker.

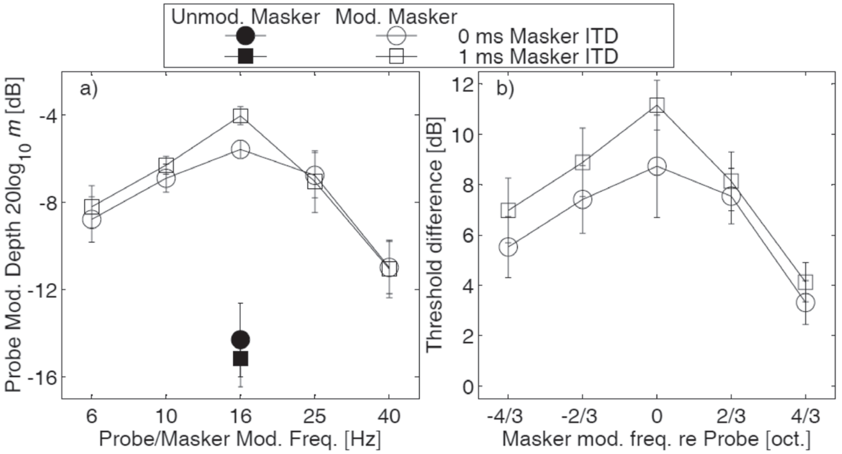

Fig. 3: Panel a: Thresholds in dB for detection of AM imposed on the probe stimulus with no AM imposed on the masker stimulus (filled symbols) and with a masker AM (open symbols), for a masker with 0-ms ITD (circles) and 1-ms ITD (squares). The probe stimulus always had a 0-ms ITD. Panel b: The difference between the thresholds measured with modulated and unmodulated maskers as a function of masker modulator center frequency and masker ITD.

The measured data show similar bandpass tuning to the data reported in the previous masked modulation detection studies, even though the stimuli used in the present study are very different from those used in previous masked modulation detection studies (e. g., Houtgast, 1989; Bacon and Grantham, 1989). Houtgast (1989) imposed the masker and probe modulators in series on a common carrier, either a broadband noise or a pure-tone, while Bacon and Grantham (1989) added the masker and probe modulators together before imposing them on a carrier. In the present study, separate carriers were used for the masker and the probe. The magnitude spectra of the masker and target carriers were identical, with only a frequency-dependent phase change related to the temporal delay. This means that the modulation spectra of the masker and probe modulators were added together, as in the study from Bacon and Grantham (1989), instead of convolved, as they are when imposed in series. These procedural differences make a quantitative comparison difficult, but, qualitatively, the prior and present studies show similar bandpass tuning in the modulation domain.

Other studies have shown an interaction between binaural listening and modulation processing. For example, in one dichotic MDI experiment (Bacon and Opie, 1994), there was a small increase in monaural modulation detection thresholds when the masker was played in the opposite ear, although much less than when both target and masker were played in the same ear. Sheft and Yost (1997) showed that ITDs in the masker slightly reduced the amount of masking in another dichotic MDI experiment. These experiments were quite different from the present study, since the target and masker modulators in those studies were imposed on carriers with a large audio-frequency difference, as opposed to the present study where they were applied to carriers with the same audio-frequency content. The MDI studies require interactions across audio-frequency channels to explain their results, while, in the present study, the modulation masking occurs within one audio-frequency channel. It is possible that the binaural interaction affects the cross-channel processing from the MDI experiments, but creates no release from masking in the within-channel masking from the present experiments.

SUMMARY AND CONCLUSIONS

Two experiments were performed using two temporally-interleaved transposed stimuli as carriers for an AM masker and for an AM probe. The lateralization experiment showed that the two stimuli could be lateralized separately using ITDs. There was a small amount of repulsion, or pushing, when the masker and probe stimuli had similar ITDs, indicating that the two stimuli were perceptually segregated.

In the masked modulation detection experiment, bandpass modulation tuning was seen in the data, but there was no release from modulation masking on the AM probe when the AM masker was lateralized with a 1-ms ITD. The measured tuning corresponds to the tuning reported in previous monaural studies of masked modulation detection.

REFERENCES

Bacon, S. P., and Grantham, D. W. (**1989**). “Modulation masking: Effects of modulation frequency, depth, and phase,” J. Acoust. Soc. Am. **85**, 2575-2580.

Bacon, S. P., and Opie, J. M. (**1994**). “Monotic and dichotic modulation detection interference in practiced and unpracticed subjects,” J. Acoust. Soc. Am. **95**, 2637-2641.

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.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (**1991**). “Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays,” J. Acoust. Soc. Am. **90**, 3077-3085.

Carlile, S., Hyams, S., and Delaney, S. (**2001**). “Systematic distortions of auditory space percep-tion following prolonged exposure to broadband noise,” J. Acoust. Soc. Am. **110**, 416-424.

Dreyer, A., and Delgutte, B. (**2006**). “Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization,” J. Neurophysiol. **96**, 2327-2341.

van der Heijden, M., and Trahiotis, C. (**1999**). “Masking with interaurally delayed stimuli: The use of ‘internal’ delays in binaural detection,” J. Acoust. Soc. Am. **105**, 388-399.

Heller, L. M., and Trahiotis, C. (**1996**). “Extents of laterality and binaural interference effects,” J. Acoust. Soc. Am. **99**, 3632-3637.

Houtgast, T. (**1989**). “Frequency selectivity in amplitude-modulation detection,” J. Acoust. Soc. Am. **85**, 1676-1680.

Kopčo, N., and Shinn-Cunningham, B. G. (**2008**). “Influences of modulation and spatial separation on detection of a masked broadband target,” J. Acoust. Soc. Am. **124**, 2236-2250.

Levitt, H. (**1971**). “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am. **49**, 467-477.

van de Par, S., and Kohlrausch, A. (**1997**). “A new approach to comparing binaural masking level differences at low and high frequencies,” J. Acoust. Soc. Am. **101**, 1671-1680.

Sheft, S., and Yost, W. A. (**1997**). “Binaural modulation detection interference,” J. Acoust. Soc. Am. **102**, 1791-1798.