Monaural envelope correlation perception for bands narrower or wider than a critical band

Emily Buss,^{a)} Joseph W. Hall III, and John H. Grose

Department of Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine,
170 Manning Drive, Chapel Hill, North Carolina 27599

(Received 2 September 2011; revised 23 October 2012; accepted 6 November 2012)

Monaural envelope correlation perception concerns the ability of listeners to discriminate stimuli based on the degree of correlation between the temporal envelopes of two or more frequency-separated bands of noise [Richards, *J. Acoust. Soc. Am.* **82**, 1621–1630 (1987)]. Previous work has examined this ability for relatively narrow bandwidths, generally 100 Hz or less. The present experiment explored a wide range of bandwidths, from 25 to 1600 Hz, which included bands narrower and wider than a critical bandwidth. Stimuli were pairs of noise bands separated by a 500-Hz-wide spectral gap centered on 2250 Hz. The magnitude spectra of the pair of comodulated bands were either identical or reflected around the midpoint of the band, and performance was assessed with and without a low-pass noise masker. Although discrimination was best for intermediate bandwidths, mean performance was above chance for all bandwidths tested. Data were similar for stimuli with identical and reflected magnitude spectra, and for stimuli with and without the low-pass masker. The one exception was particularly good performance for intermediate-bandwidth stimuli with identical spectra, for which some listeners reported hearing a tonal cue. Results indicate that listeners are flexible in selecting spectral regions upon which to base across-frequency comparisons.

© 2013 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4768887>]

PACS number(s): 43.66.Mk, 43.66.Lj, 43.66.Ki [TD]

Pages: 405–416

I. INTRODUCTION

Monaural envelope correlation perception (MECP) is a psychoacoustic paradigm that assesses the ability of listeners to discriminate between stimuli based on the degree of envelope correlation of frequency-separated bands of noise (Richards, 1987). Although there are individual differences in this task, particularly in terms of performance when the separation between noise bands is relatively wide, sensitivity to envelope correlation occurs over a wide range of band center frequencies (Richards, 1987; Moore and Emmerich, 1990) and remains robust for bands differing substantially in level (Moore and Emmerich, 1990). Similar effects have been observed in the detection of envelope phase disparities for pairs of sinusoidally amplitude-modulated tones that are separated in frequency (Strickland *et al.*, 1989). Sensitivity to MECP and envelope phase disparity has also been demonstrated for dichotic stimuli, with one band or modulated tone presented to each ear (Richards, 1989; Strickland *et al.*, 1989). This result indicates that MECP and envelope phase discrimination do not require the interaction of stimulus components from the frequency-separated bands within a single auditory filter at the periphery, although it does not rule out reliance on such a cue in some conditions.

Most MECP experiments have included, or used exclusively, a noise bandwidth of 100 Hz (Richards, 1987, 1988a,b, 1989; Moore and Emmerich, 1990; Hall and Grose, 1993). Richards (1987) investigated bandwidths that were

narrower than 100 Hz and generally found a decrease in performance with decreasing bandwidth. Moore and Emmerich (1990) noted that a 100-Hz-wide noise band has more envelope fluctuations in a given time period than a narrower band of noise. Specifically, the average fluctuation rate of a Gaussian band of noise is approximately 0.64 times its bandwidth (Rice, 1954). Moore and Emmerich (1990) suggested that the bandwidth effect could reflect an improvement in MECP with increasing numbers of envelope fluctuations available for comparison. Performance for wider bandwidths may also improve due to an increased likelihood that a particularly salient modulation may occur, as a consequence of more periods of modulation being present in the fixed-duration listening interval. This idea is consistent with the finding that the detection of envelope phase disparities is relatively good for low-rate sinusoidally amplitude-modulated tones separated in frequency (Strickland *et al.*, 1989; Green *et al.*, 1990), stimuli for which the number of modulation periods does not affect the maximum modulation depth. Whereas increasing the bandwidth of a noise stimulus improves MECP for bandwidths up to 100 Hz, the beneficial effects of further increases in bandwidth could be offset by limitations in the temporal resolution with which across-frequency envelope comparisons can be made (Strickland *et al.*, 1989). Further, when the overall stimulus level is held constant, increasing the bandwidth is associated with the introduction of high-frequency envelope components and a reduction in the power of low-frequency envelope components (Dau *et al.*, 1999).

Another possible reason for not *increasing* the noise bandwidth beyond 100 Hz in the study of MECP is related to

^{a)}Author to whom correspondence should be addressed. Electronic mail: ebuss@med.unc.edu

the frequency selectivity of the ear (Fletcher, 1940; Patterson, 1976; Glasberg and Moore, 1990). Interpretation of MECP data in terms of the processing of envelope correlation across independent peripheral frequency channels is relatively straightforward if it is assumed that the internal representations of the envelopes remain correlated at the outputs of a pair of associated auditory filters (Richards, 1987). This could occur, for example, if the spectral extent of each noise band was equal to or less than a critical band. A bandwidth of 100 Hz would appear to be a reasonable choice for MECP work, because this relatively narrow bandwidth is either close to or less than the equivalent rectangular bandwidth (ERB) for much of the audio spectrum (Glasberg and Moore, 1990).¹ Although comodulated noise bands that are much wider than the associated auditory filters can be presented, a reasonable expectation is that the envelopes at the outputs of auditory filters corresponding to the band center frequencies would be reduced in correlation, partly due to the fact that the absolute width of the auditory filter increases with increasing stimulus frequency (e.g., Fletcher, 1940; Glasberg and Moore, 1990). This view would lead to the expectation that MECP would become more difficult for noise bandwidths that are wide with respect to the auditory filter bandwidth. The present experiment tested this prediction by measuring MECP as a function of noise bandwidth, including bandwidths up to 1600 Hz, much wider than those used in previous studies.

In most previous studies of MECP, stimulus bands were generated in the frequency domain by assigning randomly selected values to the magnitude and phase components within the stimulus passband. Using this procedure, random bands are generated by the assignment of random values to each frequency bin, whereas comodulated bands are generated by the assignment of identical values to corresponding frequency bins. Comodulated bands generated in this way have identical envelopes and identical spectral profiles. In order to assess the importance of spectral coherence across bands, Richards (1988b) introduced a stimulus generation procedure that results in coherent envelopes across a pair of bands, but magnitude spectra that are mirror reflections of one another. One feature of the traditional (identical-spectrum) comodulated bands is that the temporal envelopes associated with corresponding portions of the bands are also coherent. For example, the envelopes associated with the highest 100 Hz of each of a pair of identical-spectrum comodulated 1600-Hz-wide bands are also comodulated. As a result, listeners in these conditions could base their performance on stimulus cues associated with the higher edge of each stimulus band² rather than combining information across a wider range of frequencies. In contrast, for the comodulated bands with reflected magnitude spectra, the distribution of correlated regions is also reflected in frequency; envelopes associated with the low-frequency portion of one band are correlated with those from the high-frequency portion of the other band, and vice versa. Another goal of the present experiment was to assess MECP for wide bandwidths using these two stimulus generation procedures.

One question of interest was whether listeners are flexible in selecting the spectral regions upon which to base

across-band envelope comparisons. Listeners could be inflexible in this regard if within-channel interactions contributed to performance. For example, listeners could detect envelope correlation for a pair of relatively narrow bands by monitoring the auditory filter intermediate between the low and high bands for envelope beats, which would be periodic for the identical-spectrum bands and fluctuate in rate for the reflected-spectrum bands. It is also theoretically possible that temporal fine structure could provide within-channel cues to envelope correlation for these stimuli. Dichotic conditions, where the two bands are presented to different ears, were included to assess the possible contribution of within-channel beats. Within-channel cues would also be degraded or absent for wide identical-spectrum bands, due to the fact that matched portions of the two bands would pass through different auditory filters. If MECP were based on a comparison of auditory filter outputs across frequency, as opposed to cues in an intermediate filter, then flexibility with respect to this strategy would be evident in the ability to perform the task with comparable accuracy in either the identical- or the reflected-spectrum conditions. The reflected-spectrum conditions are also interesting from the perspective of potential spectral cues. Following the logic of Richards (1988b), if performance is based on across-band comparisons of magnitude spectra, then results could be poorer for reflected-spectrum than for identical-spectrum comodulated bands.

II. METHODS

A. Listeners

Ten listeners were recruited for this study. All had previous listening experience in psychoacoustical tasks, and all had thresholds in quiet that were 20 dB hearing level or better for octave frequencies between 250 and 8000 Hz (ANSI, 2004). Initial testing with three of these listeners resulted in performance at or near chance in conditions for which other listeners performed at 80% correct or better, even after several hours of practice. These listeners were excused from further testing. An additional listener was unable to complete the experiment due to scheduling constraints. Data on the remaining six listeners are reported. Listeners L1, L2, and L4 were authors, and the remaining listeners each had 5 years or more of psychoacoustic listening experience.

B. Stimuli

Two noise bands were presented in each listening interval, a “low-frequency” band and a “high-frequency” band. For the low-frequency band the upper edge was fixed at 2000 Hz, and for the high-frequency band the lower edge was fixed at 2500 Hz. As a consequence, there was always a 500-Hz-wide spectral gap between bands, centered on 2250 Hz, regardless of band width. Seven bandwidths were investigated, spanning the range of 25–1600 Hz in one-octave steps. Each noise-band pair was presented at a level of 65 dB sound pressure level (SPL).

Bandpass noise samples were generated prior to each threshold estimation run. The low-frequency bands were generated in the frequency domain by assigning random

values to the magnitude and phase components within the noise passband, based on draws from a Rayleigh and a uniform distribution, respectively. The comodulated high-frequency bands were generated in one of two ways. In the identical-spectrum conditions, the same set of random draws—in the same sequence—was used to define the low-frequency and the high-frequency noise bands. In the reflected-spectrum conditions, the individual magnitude and phase components from the low-frequency band were assigned to sequential frequency bins of the high-frequency band in reverse order, and component phases were multiplied by -1 . Thus, the magnitude spectrum of the high-frequency band was reflected around the spectral midpoint, forming a mirror image of the low-frequency band. In random envelope conditions, independent random draws were used to define the two bands. Arrays associated with the low-frequency band, comodulated high-frequency band, and random high-frequency band were transformed into the time domain.

The time-domain stimulus arrays were loaded into different serial buffers of a real-time processor (RP2, TDT, Tucker-Davis Technologies, Alachua, FL). Each array was composed of 2^{18} points which, when played out at 24 414 Hz, resulted in a 10.7-s sample that repeated seamlessly over the course of a threshold estimation run. The amplitudes of these arrays were scaled to zero in the inter-stimulus and inter-trial intervals; stimulus bands associated with each interval were gated on for 400 ms using 30-ms raised-cosine ramps. In this way, different 400-ms segments of the 10.7-s stimulus arrays were presented in different intervals. The stimulus in each listening interval was generated by summing the low-frequency band with either the comodulated or the random high-frequency band. The stimulus was routed through a headphone buffer (HB7, TDT) and presented over circumaural headphones (Sennheiser, Old Lyme, CT, HD 265 linear).

In the course of pilot testing, several listeners reported hearing a distinct “tone-like” pitch cue associated with the comodulated bands for the 400-Hz bandwidth in the identical-spectrum condition. It is possible that this percept is associated with an auditory distortion product.³ To assess the effect of this tonal cue, a low-pass noise was introduced to mask the associated low-frequency region. The low-pass noise played continuously at 65 dB SPL, and it was generated using a sixth-order Butterworth filter with a 1000-Hz cutoff. Initially the tonal cue was thought to be restricted to the 400-Hz identical-spectrum condition, but as the experiment progressed, some listeners reported hearing analogous tones in the 100- and 200-Hz bandwidth conditions. This prompted data collection with the low-pass noise for all stimulus bandwidths, despite the fact that the masking noise spectrally overlaps with the low-frequency band for the 1600-Hz bandwidth. Listeners reported that the low-pass noise fully masked the tone-like cues audible in some conditions. No listener reported hearing a tonal cue in the reflected-spectrum conditions; nevertheless, the reflected-spectrum conditions were also completed with and without the low-pass noise masker in order to assess any effects related to the presence of masking noise, apart from its ability to mask a low-frequency distortion tone.

Although the focus of this study was *monaural* envelope correlation perception, one dichotic condition was completed in order to determine whether same-ear presentation was necessary for sensitivity to envelope correlation. This condition used the 400-Hz bandwidth, with the low-frequency band presented to the left ear and the identical-spectrum high-frequency band presented to the right ear. In the remaining conditions all stimuli were presented monaurally to the left ear.

C. Procedure

As in most previous MECP experiments, performance was quantified in terms of percent correct on fixed block trials. In the present conditions, there were 25 trials per run. On each trial there were three intervals. In one of the intervals, at random, the bands were comodulated, and in the other two intervals, the bands were random. Intervals were separated by 300-ms inter-stimulus intervals. Intervals were indicated visually with lights on a response box, and listeners made their responses by pressing buttons associated with the interval lights. Feedback was provided after each response. Five runs were completed for each condition, and all runs for a condition were completed before continuing to the next condition. Conditions were completed in quasi-random order, with the caveat that conditions with the low-pass noise tended to be completed after the conditions with no such noise. The performance estimate for a given condition was taken as the average percent correct across all five runs associated with that condition.

Listeners were given practice at all bandwidths, in monaural conditions without low-pass noise, before final data were recorded. The amount of practice varied across individual listeners, but generally included ten or more runs on each condition. Practice was extended if listeners showed continued improvement in performance, defined as a change of 15% points or more. When listener availability allowed, improvement in one condition prompted collection of replacement data at all bandwidths for the associated condition; data were fully replaced for L2, L4, and L5. In addition to this practice, at the beginning of each run listeners were given the opportunity to re-familiarize themselves with the stimuli to be tested with “reminder” trials. In the reminder trials, the comodulated bands were always presented in the second listening interval. The listener was told that she/he could listen to as many reminder trials as desired by continuing to press buttons 1 or 3. When the listener was ready to begin formal testing, button 2 was pressed to start the run of 25 trials. In total, listeners spent between 10 and 18 h total on practice and data collection.

Percent correct data were transformed prior to statistical analysis, to counteract ceiling and floor effects. The first step was to remove effects related to guessing in a three-alternative forced choice,⁴ and the second step was to apply an arcsine transform (Studebaker, 1985). While data transformed in this way were more consistent with the assumptions of parametric statistics, the same pattern of results was obtained when analyses were repeated on un-transformed percent correct data.

III. RESULTS

The results of individual listeners are plotted separately for identical-spectrum (Fig. 1) and reflected-spectrum (Fig. 2) conditions. In both cases, the mean percent correct across listeners is shown in the lower right panel. The open circles show results for the identical-spectrum or reflected-spectrum stimuli presented in quiet, closed circles show data for conditions in which the low-pass noise was present, and the horizontal dashed line indicates chance performance. The triangles in Fig. 1 show data for the dichotic condition. Dotted lines in Fig. 2 show thresholds from the associated identical-spectrum conditions with low-pass noise, replotted from Fig. 1 to facilitate comparison of the two methods for generating comodulated noise bands. The identical-spectrum conditions with low-pass noise were chosen for comparison due to the fact that performance was likely unaffected by the presence of the tonal cue heard by some listeners with identical-spectrum stimuli in quiet. For individual data, error bars show the range of estimates obtained across the five runs. For mean data, error bars show the range of mean estimates obtained by individual listeners. Although there were individual differences, all listeners achieved performance levels above chance for most conditions.

Across listeners and stimulus conditions, the best performance appeared to occur for intermediate bandwidths of 100–400 Hz. In general, performance was similar with and without low-pass noise, and for the identical-spectrum and reflected-spectrum stimuli. The one possible exception to

this trend was the relatively good performance in the identical-spectrum condition without low-pass noise for the 400-Hz bandwidth condition and, to a lesser extent, the 200-Hz condition in some listeners. This trend is evident in data for L2 and L3, both listeners who reported hearing a tonal percept in the absence of the low-pass noise. Initial data for L5 (not shown) exhibited a similar trend, with 100% correct for the 100- and 200-Hz identical-spectrum conditions in the absence of low-pass noise. Data were subsequently replaced for this listener, due to improvement in performance for the wide bandwidths. In these new data, performance dropped to 69% and 78% for 100- and 200-Hz bandwidths, respectively. It is possible that intermediate exposure to the low-pass noise conditions caused this listener to focus on stimulus cues other than the tonal percept.

A repeated-measures analysis of variance (ANOVA) was performed to evaluate trends in the group data. There were two levels of condition (identical- and reflected-spectrum), two levels of masker (none and low-pass), and seven levels of bandwidth (25, 50, 100, 200, 400, 800, and 1600 Hz). This analysis resulted in a main effect of bandwidth ($F_{6,30} = 30.057$, $p < 0.001$), and no main effect of either condition ($F_{1,5} = 0.616$, $p = 0.468$) or masker ($F_{1,5} = 0.220$, $p = 0.659$). None of the interactions reached significance ($p > 0.05$). Notably, the three-way interaction approached but did not attain significance ($F_{6,30} = 2.304$, $p = 0.060$). The main effect of bandwidth was further evaluated with contrasts. The quadratic contrast with bandwidth

FIG. 1. The open circles show mean percent correct data as a function of stimulus bandwidth for the identical-spectrum conditions. The closed circles show results in the presence of a low-pass noise. The triangles show data for the dichotic 400-Hz bandwidth condition, in which the low- and high-frequency identical-spectrum bands were presented to the left and right ears, respectively. Individual listeners' data are shown in different panels, and the mean across listeners appears at the lower right. Error bars show the range of values obtained in the five runs for each individual or the range of mean values across individuals, and the dashed horizontal line indicates chance performance.

FIG. 2. Results for the reflected-spectrum conditions are shown, following the plotting conventions of Fig. 1. The dotted line indicates data for the identical-spectrum stimulus in low-pass noise, replotted from Fig. 1 to aid comparison across conditions.

was significant ($F_{1,5} = 134.636$, $p < 0.001$), but the linear contrast was not ($F_{1,5} = 4.338$, $p = 0.092$). This finding lends support for the conclusion that performance was better at intermediate bandwidths than at either the low or high bandwidths. The finding of reduced sensitivity at the lowest bandwidths agrees with previous studies that compared performance for a 100-Hz bandwidth to performance for narrower bandwidths (Richards, 1987; Moore and Emmerich, 1990). The finding of reduced sensitivity at the highest bandwidths tested is novel.

One possible reason for the absence of a significant three-way interaction lies in individual differences in the ability to hear a tonal cue. In those listeners who reported hearing this cue (L2, L3, and L5), low-pass noise tended to reduce sensitivity for intermediate bandwidths in the identical-spectrum condition. In those listeners who were unable to hear this cue (L4 and L6), low-pass noise had little or no effect. While L1 reported hearing a tonal cue in quiet, his performance was near ceiling in the identical-spectrum condition regardless of the presence of low-pass masking noise, perhaps due to a robust ability to capitalize on the remaining cues. In most conditions, listeners reported that the cue for comodulation was a “rough,” “scratchy,” “crackly,” “grindy,” or “static-like” quality. While the percept differed somewhat in detail across bandwidths, it was very similar across identical-spectrum and reflected-spectrum conditions with and without low-pass noise. In cases where a tonal percept was audible, the introduction of low-pass noise effectively

masked that percept and did not have an obvious negative effect on the ability to use the other, more common cue. This result indicates that the presence of the continuous low-pass noise itself probably does not interfere with MECP.

The dichotic presentation of 400-Hz-wide identical-spectrum bands was consistently associated with performance above chance. Across listeners, performance in the identical-spectrum 400-Hz bandwidth condition was significantly poorer for dichotic than the monaural presentation without the low-pass noise ($t_5 = 4.830$, $p = 0.005$), with a mean difference of 20.4%. As for monaural presentation, listeners reported a percept related to roughness in the dichotic conditions when the bands were comodulated. Performance in the dichotic condition was not significantly different from the 400-Hz identical-spectrum condition with low-pass noise ($t_5 = 1.398$, $p = 0.221$), suggesting that the dominant effect of dichotic presentation could be the elimination of the tonal cue.

IV. DISCUSSION

A. Overview of effects related to bandwidth and noise condition

The primary goal of the present experiment was to evaluate MECP for a range of bandwidths, including bandwidths spanning multiple auditory filters. In general, thresholds were best for intermediate bandwidths, with poorer performance for lower and higher bandwidths. The bandwidths

tested here span the range of approximately 0.1–11.8 ERBs, such that some stimuli were subcritical and others were wider than an ERB. Results for the three lowest bandwidths are broadly consistent with previous reports showing greater sensitivity in MECP for the 100-Hz bandwidth than narrower bandwidths. This effect is striking in some listeners (e.g., L2) and absent in others (e.g., L3). Similar individual differences were observed in previous MECP studies comparing performance for 100-Hz and narrower bandwidths (Richards, 1987; Moore and Emmerich, 1990). A novel finding of the present study is that listeners were able to reliably discriminate comodulated from random stimulus bands for bandwidths considerably wider than an ERB. For example, in the 1600-Hz conditions the low- and high-frequency bands spanned 11.8 and 4.3 ERBs, respectively. Although MECP was achieved for these stimuli, performance was not as good as for intermediate bandwidths. Possible reasons for this are discussed in the following.

Sensitivity to envelope correlation was similar for the identical-spectrum and reflected-spectrum stimuli, and performance with and without the continuous low-pass noise was likewise similar. The exception to both of these trends occurred for the intermediate bandwidths of the identical-spectrum stimulus, the conditions for which some listeners reported hearing a tonal cue. In some listeners, percent correct in the 400-Hz identical-spectrum condition was reduced by the introduction of low-pass noise, and performance with the low-pass noise masker closely matched that in the reflected-spectrum condition, with or without low-pass noise. A less pronounced trend for better performance in the absence of low-pass noise for the 200-Hz bandwidth identical-spectrum condition was also seen in individual data, as well as practice data for the 100-Hz bandwidth in one listener. A parsimonious explanation for this result is that some listeners were able to make use of a tonal cue in the mid-bandwidth identical-spectrum conditions, but that this cue was unavailable in the reflected-spectrum and low-pass noise conditions.

A priori, it was hypothesized that the reflected-spectrum stimuli would be associated with poorer performance than the identical-spectrum conditions if listeners were inflexible in their ability to selectively compare envelopes across spectral regions, or in their ability to compare magnitude spectra of the two bands. The data are not consistent with this view, however. There were also *a priori* reasons to expect that the low-pass masker might decrease MECP performance, apart from considerations related to the masking of a tonal cue. This could occur if the ability of the auditory system to include only specific frequency regions in MECP analysis were limited. For example, if the low-pass noise could not be excluded from the analysis, then performance would deteriorate because the low-pass noise was uncorrelated with respect to the target noise bands. Another possibility is that the auditory system might be able to exclude the low-pass noise from the envelope information included in the MECP analysis, but only at a cost; the increased task complexity involving perceptual segregation of the low-pass noise from the target noise bands could reduce overall performance. However, neither of these *a priori* possibilities was consistent

with the findings, as the only clear effect of the low-pass noise was the masking of a tonal percept that occurred in the identical-spectrum condition. It therefore seems reasonable to conclude that across-frequency processing in MECP was not influenced either by the presence of low-pass noise or reflection of the spectrum.

Two conditions that warrant further consideration are the 1600-Hz conditions with low-pass noise. In both the identical-spectrum and reflected-spectrum conditions, inclusion of the low-pass noise had no apparent effect on performance despite the introduction of substantial energetic masking in the region of the low-frequency band. The 1000-Hz low-pass masker and the low-frequency 1600-Hz-wide band overlapped in the region of 400–1000 Hz, where the signal-to-noise ratio was –5 dB. In the comodulation condition, the Pearson product correlation between the stimulus envelope in this region of the low-frequency band and the associated high-frequency stimulus was reduced to $r = 0.22$ by inclusion of the low-pass masker. This degree of decorrelation has been shown to adversely affect MECP, reducing sensitivity to near chance under some conditions (Richards, 1987). The question this observation raises is why the inclusion of the low-pass masker did not hurt performance for the 1600-Hz bandwidth.

One possible explanation for the apparent irrelevance of the low-pass masker is that the region between 400 and 1000 Hz did not contribute to performance even in the conditions without the low-pass masker. This possibility gains some support from the observation that MECP tends to be poor for low-frequency stimuli. For example, Richards (1987) observed that MECP for 100-Hz-wide bands as a function of center frequency was best for 4000 and 2500 Hz, reduced for 1000 Hz, and at chance for 350 Hz. An effect of carrier frequency has also been observed for the detection of envelope phase disparity with sinusoidally amplitude-modulated tones, where sensitivity was best in the region of 4000 Hz (Green *et al.*, 1990). The source of this carrier frequency effect is unknown, but it is possible that representation of temporal fine structure in the auditory system reduces the quality of envelope information (e.g., Rhode and Greenberg, 1994). A subtler cue at low than high frequencies is broadly consistent with the idea that individual differences in MECP are particularly pronounced at 350 Hz (Moore and Emmerich, 1990). This explanation is somewhat inconsistent with the finding of a robust comodulation masking release at low frequencies (e.g., Schooneveldt and Moore, 1987), however, in that comodulation masking release (like MECP) is thought to rely on the use of envelope cues. Regardless of the reason behind the frequency effect in MECP, if frequencies below 1000 Hz do not contribute significantly to MECP in the present 1600-Hz bandwidth conditions, then energetic masking of these stimulus components would not necessarily hurt performance.

B. Consideration of possible within-channel effects

The dichotic, 400-Hz bandwidth condition explored here is of interest because it eliminates the possibility of a peripheral, within-channel beating cue. Note that this

condition has a possible confound because dichotic presentation also tends to perceptually segregate the stimulus components presented to each ear. Such segregation has been argued to pose a difficulty for across-frequency envelope comparisons in other paradigms (e.g., Buss and Hall, 2008). No such difficulties were evident in the present results, however. Here, performance in the dichotic condition was similar to that obtained for the corresponding identical-spectrum condition where a low-pass noise was presented to mask the tonal cue. This similarity between the results for the 400-Hz bandwidth dichotic and masked monaural conditions is consistent with an interpretation that listeners did not rely on within-channel cues in the monaural condition, although it does not rule out a possible contribution of within-channel cues for stimulus bandwidths other than 400 Hz. The dichotic stimulus data are qualitatively similar to the supplemental MECP data reported in Hall and Grose (1993).

A second observation that speaks against the use of within-channel cues is the absence of an effect of stimulus generation method. For narrowband, comodulated stimuli, the envelope at the output of the auditory filter intermediate between the bands could provide a cue to band correlation. For the identical-spectrum condition, correlation is associated with a pronounced modulation corresponding to the separation between bands. For the reflected-spectrum condition, correlation is associated with pronounced modulation that changes in rate over time, such that the envelope difference between correlated and random bands may not be as pronounced for the reflected-spectrum as for the identical-spectrum conditions. Broadening the stimulus bands results in a reduction or elimination in the periodic envelope beats associated with the identical-spectrum stimulus. However, the variable-rate envelope fluctuation associated with the reflected-spectrum condition persists. If within-channel cues were providing an important cue to envelope correlation, then the pattern of data as a function of bandwidth would be expected to differ for the identical- and reflected-spectrum conditions. No such differences were observed.

C. Auditory processes that may underlie MECP for relatively wide noise bandwidths

We now consider two general explanations for the robust MECP sensitivity observed in this study for relatively wide noise bandwidths. One is based on envelope comparisons associated with auditory filter outputs corresponding to the best-matched frequency regions of the low and high noise bands. This will be referred to as “narrowband analysis.” The other is based upon envelope comparisons associated with central filters that combine information across frequency. This will be referred to as “wideband analysis.” Before considering these processes in detail, it is instructive to consider the peripheral encoding of the stimulus.

The top panel of Fig. 3 shows a pair of 1600-Hz-wide comodulated, identical-spectrum bands in the frequency domain. The bottom panel shows a set of six roex filters, generated using the procedures described by Moore and Glasberg (1987) and uniformly distributed across each of the two stimulus bands. These filters are designed to represent the

limits of spectral resolution of the peripheral auditory system. In this example, envelopes would be most highly correlated at the outputs of filter pairs 1 and 4, 2 and 5, and 3 and 6. For a reflected-spectrum stimulus (not shown), maximum correlations would reflect the rotated spectral structure of the bands, with largest values for filter pairs 1 and 6, 2 and 5, and 3 and 4. The correlations between pairs of auditory filter outputs would be positive for both identical- and reflected-spectra stimuli, but they would fall short of $r = 1$. This is due to the fact that auditory filters are not uniform across frequency, most notably increasing in width with frequency.

1. MECP based upon comparisons between auditory filter outputs

When each stimulus band is wide with respect to the associated auditory filters, the auditory system could detect envelope correlation by comparing the outputs of specific pairs of auditory filters. In this scheme, the final decision cue could be based on either a single comparison or a conglomeration of multiple such comparisons. A simulation was performed to illustrate the basic features of envelope comparisons across matched auditory filters for the 800-Hz bandwidth stimuli. A bank of roex filters (Moore *et al.*, 1987) was generated, with 10-Hz separation between neighboring filters across the spectral range of the stimulus. Additional off-frequency filters were included, spanning one ERB above and below the stimulus bands. This was designed to capture off-frequency listening. Stimuli were generated, as described previously, and passed through the roex filterbank. The Hilbert envelopes at the outputs of each filter were computed. This process was repeated 100 times, and the median cross correlation for each pair of filter outputs was recorded.⁵ This approach captures level and frequency effects of the auditory filter, but does not incorporate phase effects, amplitude compression, or dynamic level effects at the periphery. While more physiologically realistic models exist (e.g., Patterson *et al.*, 2003), roex filters were preferred in the present application due to their simplicity.

Figure 4 shows the median correlations generated by this simulation, with lighter shading indicating higher correlation, as defined in the legend, and dashed lines above the diagonal indicating the frequency regions associated with the low- and high-frequency bands. Correlations are shown

FIG. 3. Illustration of stimuli and roex filters. (A) A 1600-Hz identical-spectrum comodulated stimulus and (B) a selection of three pairs of roex filters with matched center frequencies in the two bands.

FIG. 4. Envelope correlations at the outputs of roex filters for a pair of 800-Hz-wide noise bands, plotted as a function of filter center frequency in kHz. Shading indicates the magnitude of the cross correlation, as defined in the legend. Panels show results for the identical-spectrum (left), reflected-spectrum (middle), and random-spectrum (right) stimuli. Dashed lines indicate the frequency regions associated with the low- and high-frequency bands.

for three stimulus conditions: Identical-spectrum (left), reflected-spectrum (center), and random (right) bands. In all three panels there is a region of white along the diagonal, indicating high correlations between envelopes at the outputs of neighboring filters. This is easiest to see in the far right panel, showing results for the random noise bands. The envelope correlation between associated spectral regions of coherently modulating bands is indicated by light shading parallel to the diagonal for the identical-spectrum stimulus (left panel) and orthogonal to the diagonal for the reflected-spectrum stimulus (middle panel). For the identical-spectrum condition, the correlation is relatively high for all matched pairs of filters, with a trend for greater correlation across filter outputs associated with the upper edges of both bands. In contrast, correlations associated with the reflected-spectrum condition are larger for filter outputs near the inner edges of each band than for the outer edges. This reflects the larger mismatch in auditory filter widths at the frequency extremes.

If the auditory system performs across-frequency narrowband analyses in the MECF task, then one possibility is that performance is based on the outputs of auditory filter pairs associated with the highest envelope correlations. Another possibility is that MECF for relatively wideband stimuli is based at least in part upon a combination of information across the frequency-matched pairs of auditory filters. Such a process is consistent with a modulation filterbank model. In this approach, auditory filter outputs are passed through modulation filterbanks, and the results are optimally combined (Dau *et al.*, 1996; Dau *et al.*, 1997b). This class of model can discriminate random from comodulated bands (Piechowiak *et al.*, 2007).

One approach to evaluating a narrowband analysis is to consider how performance changes with stimulus bandwidth (see Figs. 1 and 2). In the identical-spectrum conditions, one interesting feature is that as bandwidth increases, the matched Fourier components of the comodulated bands become increasingly separated in frequency. This is because the lower components of the high-frequency band correspond to the lower components of the low-frequency band: The lower components of the high-frequency band remain fixed near 2500 Hz, but the lower components of the

low-frequency band progressively decrease in frequency as the bandwidth increases. One ramification of this is that as bandwidth is increased, the mismatch in auditory filter bandwidth for corresponding filters increases. The increasing auditory filter bandwidth mismatch with increasing stimulus bandwidth for the identical-spectrum stimulus could contribute to the downturn in performance as bandwidth increased beyond 400 or 800 Hz (see Fig. 1).

Although increases in stimulus bandwidth also result in the introduction of stimulus components associated with progressively larger mismatches in auditory filter bandwidth in the reflected-spectrum conditions, there are important differences in this effect when contrasted with the identical-spectrum conditions. In the reflected-spectrum conditions, each increase in bandwidth introduces “new” stimulus components at the low edge of the low-frequency band and the high edge of the high-frequency band. As a consequence, the bandwidths of corresponding auditory filters are quite similar for filter pairs near 2000 and 2500 Hz, present in both narrow and wide stimuli. Increasing the stimulus bandwidth introduces additional filter pairs with increasingly disparate bandwidths. This can be seen in Fig. 4, where correlation for the reflected spectrum is relatively low for the filter pairs including the low edge of the low-frequency band and the high edge of the high-frequency band.

Increasing mismatches in the bandwidths of paired auditory filters could contribute to the reduced sensitivity for the widest stimulus bandwidths, but this explanation is not wholly consistent with the data. For one thing, this factor would predict a differential bandwidth effect for the identical- and reflected-spectrum condition, which did not occur. Further, it is unclear how increasing the bandwidth in reflected-spectrum conditions could hurt performance. Although an increase in bandwidth in the reflected-spectrum conditions adds additional energy that is associated with auditory filters that are progressively less well matched in bandwidth, the noise components associated with the narrower bandwidths remain predominantly comodulated. Therefore, accounts based upon the combination of information from the auditory filters associated with the best matches would predict improvement in performance as the bandwidth increased, but little or no decrease in performance for bandwidths

above 400–800 Hz. One possibility is that the auditory system is not able to use the best information available, but instead weights the information across bandwidth in a suboptimal way. This is not wholly consistent with the low-pass noise data, however, where the introduction of the low-pass masker had no substantial effect on performance except in cases where a distortion tone was being masked. This apparent discrepancy could be mitigated by considering the greater ease of segregating out the low-pass noise, due to its continuous presentation and spectral separation from the target stimulus. These segregation cues would not be available to the listener in selecting a subset of auditory channels to attend to in the wideband, comodulated stimulus.

2. MECP based upon comparisons involving relatively wide central filters

Some previous work on envelope processing for relatively wide noise bandwidths has suggested the possibility that performance may be based on the output of a broad “predetection filter” that subsumes a wide range of peripherally based auditory filters. For example, Viemeister (1979) and Forrest and Green (1987) found that this sort of wideband process, followed by a low-pass filter, was successful in accounting for the detection of modulation of wideband noise carriers. This approach is considerably less accurate in predicting modulation detection with narrowband stimuli, however; Dau *et al.* (1997a) showed that the modulation filterbank is better able to capture modulation detection with narrowband stimuli than a model based on low-pass filtering. Recently, Berg (2007) suggested that a wideband analysis could account for some of the data patterns in a profile analysis task (Green, 1988), involving the detection of an increment in the center component of a spectral profile that was roved in level. Berg found results consistent with the idea that some listeners relied on temporal envelope information that was derived by combining information across many auditory filters, spanning a range of up to 715 Hz in the region of 1000 Hz. Specifically, Berg (2007) proposed that, “Arbitrarily broad filters could be constructed by combining the input of many primary fibers, even those emanating from resolved locations on the basilar membrane.” Applying this approach to the current stimulus set, the auditory system could combine information from the low frequencies to reconstruct the low band and information from the high frequencies to reconstruct the high band. Envelopes of the two reconstructed bands could then be extracted and compared.

It is possible that neurons at a very peripheral stage could underlie this kind of wideband analysis. For example, many onset units in the cochlear nucleus receive inputs from multiple primary units associated with a relatively broad range along the basilar membrane (Golding *et al.*, 1995; Jiang *et al.*, 1996; Pressnitzer *et al.*, 2001), and such units demonstrate responses related to the envelope of a sound (Frisina *et al.*, 1990; Rhode and Greenberg, 1994). It is possible that higher levels of auditory analysis could support selective processing of the outputs of these onset units. This would allow flexibility in the analysis of information associated with specific frequency regions, including some regions

(e.g., those associated with the best stimulus cues) and omitting others (e.g., those associated with the low-pass masker). Results obtained with low-pass noise are in agreement with the idea that there may be flexibility in the bandwidth used in the process: There was little or no reduction in sensitivity with the introduction of the low-pass noise, suggesting that the added noise could be largely excluded from the process underlying performance.

It is of interest to consider whether a wideband analysis is compatible with the downturn in performance observed for the widest bandwidths tested in the present study. One possibility is that this effect is related to the increase in fluctuation frequency as noise bandwidth is increased. Such an interpretation would be consistent with previous studies showing reduced sensitivity to temporal envelope at relatively high fluctuation frequencies. For example, although Viemeister’s (1979) temporal modulation transfer (TMTF) study using a noise carrier showed sensitivity to fluctuation up to a modulation frequency of approximately 1000 Hz, performance began to roll off at approximately 64 Hz. The ability to discriminate modulation rate with a noise-band carrier also falls off above about 70 Hz (Patterson *et al.*, 1978), a finding that has been attributed to the reduced detectability of envelope fluctuation at high rates. The TMTF function for pure-tone carriers begins to roll off at approximately 150 Hz (Kohlrausch *et al.*, 2000), and the ability to detect envelope phase differences in sinusoidally amplitude-modulated tones also falls off above 100 Hz (Green *et al.*, 1990). These results are consistent with the idea that the downturn in performance found here for the wider bandwidths may reflect a general decrease in sensitivity to stimulus fluctuation for relatively high modulation rates. Interpretation of the present findings at the widest bandwidths in terms of reduced ability to follow fast envelope fluctuation is complicated somewhat by the fact that the noise stimuli had complex modulation spectra. Thus, even though the reconstructed stimulus envelope may have had a relatively high average fluctuation rate, envelope fluctuation rates below and above the average rate were also present; it is unclear why the task could not be performed based on the well-represented low-frequency envelope features. It is possible that the reduction in low-frequency envelope power with increasing bandwidth for a fixed-level stimulus reduces the utility of low-frequency envelope features (Dau *et al.*, 1999).

An alternative hypothesis is that performance with the wide stimulus bands is limited by the reduced access to envelope information for stimuli in low-frequency regions. Recall that masking data for the 1600-Hz bandwidth indicate no reduction in performance when the 400–1000 Hz portions are masked. It is possible that this effect is related to the previously discussed finding by Richards (1987) that MECP performance deteriorated for noise band frequencies near 1000 Hz and lower. If this region of the low-frequency band did not contribute significantly to the auditory representation of the envelope of the lower band, but the corresponding frequencies did contribute to the envelope of the high-frequency band, this discrepancy could have reduced the quality of the overall decision cue. Thus, it is possible that reduced contribution from the lower frequencies of the

low-frequency band could account for both the lack of a low-pass masker effect and the downturn in performance for the wider bandwidths.

3. Supplemental data: Randomizing correlated spectral regions within a block of trials

At the outset of this experiment, it was anticipated that MECP with stimuli that are wider than a critical band could be difficult, in that it could require detailed knowledge about the pairs of auditory filter outputs upon which to base performance. Further, it was expected that changing this correspondence, with identical- and reflected-spectrum stimuli, could cause additional difficulties, in that listeners would need to use different listening strategies in these two conditions. The subjective reports of the listeners in this experiment belied those expectations, however. Listeners reported that the cues in identical- and reflected-spectrum conditions were the same.

During the review process, it was suggested that if different strategies were applied when performing MECP in the two stimulus conditions, then randomly interspersing identical- and reflected-spectrum stimuli should reduce sensitivity. Supplemental data were collected using this stimulus manipulation. One listener from the main experiment (L3) was unavailable for testing, and so was replaced. Performance was assessed using methods identical to those described previously. The stimuli were 800-Hz-wide bands of noise, and the target interval contained comodulated bands. The 800-Hz bandwidth was chosen because it spans multiple auditory filters, and performance was relatively good in most listeners in the primary dataset (~80% correct), such that a reduction in performance would be evident. In one condition the comodulated stimuli were identical-spectrum bands, in another they were reflected-spectrum bands, and in a third they were either identical- or reflected-spectrum bands, selected at random with equal probability on each trial (mixed). While the first two conditions were identical to those collected in the main experiment, it was decided to take new data in these conditions due to the fact that more than six months had passed since the initial data had been collected. Data were collected blocked by condition, with conditions visited in random order.

The supplemental data are shown in Fig. 5. Individual data are indicated with open circles, ordered on the abscissa by listener number, and mean data are shown with closed circles. Mean performance was quite similar across conditions, with values of 74%–77% correct. A repeated-measures ANOVA with three levels of masker condition (identical spectrum, reflected spectrum, and mixed) confirmed this impression ($F_{2,10}=0.516$, $p=0.612$). This result is consistent with the subjective impression that the same cue underlies performance in the identical- and reflected-spectrum conditions, in that the predictability of stimulus type did not affect performance. This result seems to be more consistent with a wideband than a narrowband process. A narrowband process, as it is formulated here, would rely on different listening strategies or cues in the identical- and reflected-spectrum conditions. While it is possible that the listener

FIG. 5. Supplemental data are shown, indicating the effect of randomly interspersing identical- and random-spectrum conditions. Individual listeners' percent correct scores are plotted with open circles in each of three conditions, with points ordered according to listener number. The mean across listeners is indicated with closed circles.

could use different strategies or cues when assessing stimuli in different intervals, it is reasonable to expect that there would be a cost associated with this; there is no evidence of such a cost in the data. Further, the apparent lack of high-level cognitive listening strategies specific to the two stimulus conditions lends some credence to the idea that the combination of information across frequency could occur relatively early in auditory processing.

V. CONCLUSIONS

The results of this study support the following conclusions:

- (1) For the frequency region examined here, comodulated and random bands of noise can be discriminated for bandwidths ranging from 25 to 1600 Hz.
- (2) While tonal pitch cues may be present in a subset of conditions, MECP with 400-Hz-wide bands can be performed in the presence of low-pass noise and for dichotic presentation, indicating that this task does not rely on the availability of this tonal cue or within-channel cues (e.g., those based on beating).
- (3) The results of conditions using identical-spectrum and reflected-spectrum stimuli suggest that performance is based on a relatively sophisticated across-channel process. This may be related to previous data suggesting an ability to group stimulus components based on coherent amplitude modulation (Bregman *et al.*, 1985; Hall and Grose, 1990).
- (4) It is not presently clear what type of auditory analysis underlies performance when noise bands are wide relative to the associated ERBs. One possibility is that MECP is based upon the correlation of envelopes at the outputs of paired peripheral auditory filters. Another possibility is that performance is based upon the envelopes at the outputs of wider filters that are constructed from signals carried by primary neural fibers associated with a wider range of frequencies. Further work is required to

determine the nature of the cue or cues supporting MECP for bands wider than an ERB.

ACKNOWLEDGMENTS

This research was supported by NIDCD Grant Nos. R01-000418 (J.W.H.) and R01-DC007391 (E.B.). This work was improved by comments from Torsten Dau, Steven van de Par, Heather Porter, and two anonymous reviewers.

¹This argument should be regarded as a first approximation only, and is complicated by the fact that the correlation of the envelopes at the outputs of auditory filters could be reduced by across-frequency differences in filter shape and phase response.

²Comparing envelopes at the output of a pair of auditory filters separated in frequency does not necessarily imply that the listeners can “hear out” these cues, in the sense of segregating these stimulus components from the remaining components. As in cases where off-frequency listening has been shown to improve detection thresholds (Johnson-Davies and Patterson, 1979), a listener might base detection on cues that are features of a more general percept.

³Although tangential to the main issues studied here, it is of interest to consider how a tonal distortion product might arise from stimulation with a pair of 400-Hz-wide bands of identical-spectrum noise, separated by 500 Hz. One listener was able to match the pitch of the percept to a 900-Hz pure tone. When a low-level tone close in frequency to 900 Hz was added, the listener reported beats with a periodicity equal to the frequency difference between the added tone and 900 Hz. For the stimulus composed of 400-Hz-wide noise bands (1600–2000 and 2500–2900 Hz), the frequency of 900 Hz corresponds to the difference frequency between the corresponding Fourier components of the two noise bands. It is perhaps worth noting that all corresponding pairs of Fourier components in this stimulus were matched in magnitude, and that this may have resulted in relatively large difference tones. Although the 900-Hz difference tones would be expected to have random phase, the sum of many such components may have risen above a noise floor composed, in part, of other distortion products. The failure to hear a tone for the 800-Hz bandwidth could be due to the fact that the expected difference tone of 1300 Hz would be masked by the lower edge of the low-frequency band (extending from 1200 to 2000 Hz). Although attempts to measure distortion tones at the output of the headphones failed, we cannot rule out the possibility that these distortions were introduced in hardware.

⁴Percent correct values (x) were adjusted to remove effects related to guessing in a 3-alternative forced choice as follows: $y = [x - (100/3)](3/2)$. In one condition for one listener, this resulted in a negative value; that value was set to zero.

⁵While previous work on MECP has used the Pearson correlation (Richards, 1987), the envelope cross correlation was used here based on work by van de Par and Kohlrausch (1998) showing that cross correlation provides a better fit to many psychoacoustical data. Repeating this analysis using Pearson correlations does not change the pattern of results shown in Fig. 4.

ANSI (2004). S3.6, *American National Standard Specification for Audio-meters* (American National Standards Institute, New York).

Berg, B. G. (2007). “Estimating the transition bandwidth between two auditory processes: Evidence for broadband auditory filters,” *J. Acoust. Soc. Am.* **121**, 3639–3645.

Bregman, A. S., Abramson, J., Doehring, P., and Darwin, C. J. (1985). “Spectral integration based on common amplitude modulation,” *Percept. Psychophys.* **37**, 483–493.

Buss, E., and Hall, J. W. (2008). “Factors contributing to comodulation masking release with dichotic maskers,” *J. Acoust. Soc. Am.* **124**, 1905–1908.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997a). “Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,” *J. Acoust. Soc. Am.* **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (1997b). “Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,” *J. Acoust. Soc. Am.* **102**, 2906–2919.

Dau, T., Puschel, D., and Kohlrausch, A. (1996). “A quantitative model of the ‘effective’ signal processing in the auditory system. I. Model structure,” *J. Acoust. Soc. Am.* **99**, 3615–3622.

Dau, T., Verhey, J., and Kohlrausch, A. (1999). “Intrinsic envelope fluctuations and modulation-detection thresholds for narrow-band noise carriers,” *J. Acoust. Soc. Am.* **106**, 2752–2760.

Fletcher, H. (1940). “Auditory patterns,” *Rev. Mod. Phys.* **12**, 47–65.

Forrest, T. G., and Green, D. M. (1987). “Detection of partially filled gaps in noise and the temporal modulation transfer function,” *J. Acoust. Soc. Am.* **82**, 1933–1943.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990). “Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement,” *Hear. Res.* **44**, 99–122.

Glasberg, B. R., and Moore, B. C. J. (1990). “Derivation of auditory filter shapes from notched-noise data,” *Hear. Res.* **47**, 103–138.

Golding, N. L., Robertson, D., and Oertel, D. (1995). “Recordings from slices indicate that octopus cells of the cochlear nucleus detect coincident firing of auditory nerve fibers with temporal precision,” *J. Neurosci.* **15**, 3138–3153.

Green, D. M. (1988). *Profile Analysis* (Oxford University Press, New York).

Green, D. M., Richards, V. M., and Onsan, Z. A. (1990). “Sensitivity to envelope coherence,” *J. Acoust. Soc. Am.* **87**, 323–329.

Hall, J. W., and Grose, J. H. (1990). “Comodulation masking release and auditory grouping,” *J. Acoust. Soc. Am.* **88**, 119–125.

Hall, J. W. I., and Grose, J. H. (1993). “Monaural envelope correlation perception in listeners with normal hearing and cochlear impairment,” *J. Speech Hear. Res.* **36**, 1306–1314.

Jiang, D., Palmer, A. R., and Winter, I. M. (1996). “Frequency extent of two-tone facilitation in onset units in the ventral cochlear nucleus,” *J. Neurophysiol.* **75**, 380–395.

Johnson-Davies, D., and Patterson, R. (1979). “Psychophysical tuning curves: Restricting the listening band to the signal region,” *J. Acoust. Soc. Am.* **65**, 765–770.

Kohlrausch, A., Fassel, R., and Dau, T. (2000). “The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers,” *J. Acoust. Soc. Am.* **108**, 723–734.

Moore, B. C., and Emmerich, D. S. (1990). “Monaural envelope correlation perception, revisited: Effects of bandwidth, frequency separation, duration, and relative level of the noise bands,” *J. Acoust. Soc. Am.* **87**, 2628–2633.

Moore, B. C., and Glasberg, B. R. (1987). “Formulae describing frequency selectivity as a function of frequency and level, and their use in calculating excitation patterns,” *Hear. Res.* **28**, 209–225.

Moore, B. C. J., Poon, P. W., Bacon, S. P., and Glasberg, B. R. (1987). “The temporal course of masking and the auditory filter shape,” *J. Acoust. Soc. Am.* **81**, 1873–1880.

Patterson, R. D. (1976). “Auditory filter shapes derived with noise stimuli,” *J. Acoust. Soc. Am.* **59**, 640–654.

Patterson, R. D., Johnson-Davies, D., and Milroy, R. (1978). “Amplitude-modulated noise: The detection of modulation versus the detection of modulation rate,” *J. Acoust. Soc. Am.* **63**, 1904–1911.

Patterson, R. D., Unoki, M., and Irino, T. (2003). “Extending the domain of center frequencies for the compressive gammachirp auditory filter,” *J. Acoust. Soc. Am.* **114**, 1529–1542.

Piechowiak, T., Ewert, S. D., and Dau, T. (2007). “Modeling comodulation masking release using an equalization-cancellation mechanism,” *J. Acoust. Soc. Am.* **121**, 2111–2126.

Pressnitzer, D., Meddis, R., Delahaye, R., and Winter, I. M. (2001). “Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus,” *J. Neurosci.* **21**, 6377–6386.

Rhode, W. S., and Greenberg, S. (1994). “Encoding of amplitude modulation in the cochlear nucleus of the cat,” *J. Neurophysiol.* **71**, 1797–1825.

Rice, S. O. (1954). “Mathematical analysis of random noise,” in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York), pp. 133–294.

Richards, V. M. (1987). “Monaural envelope correlation perception,” *J. Acoust. Soc. Am.* **82**, 1621–1630.

Richards, V. M. (1988a). “Aspects of monaural synchrony detection,” in *Basic Issues in Hearing: Proceedings of the 8th International Symposium on Hearing* (Academic, New York), pp. 317–322.

Richards, V. M. (1988b). “Components of monaural envelope correlation perception,” *Hear. Res.* **35**, 47–57.

Richards, V. M. (1989). “Comparing monotic, diotic, and dichotic presentation modes in synchrony detection,” *J. Acoust. Soc. Am.* **85**, S143.

- Schooneveldt, G. P., and Moore, B. C. J. (1987). "Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band," *J. Acoust. Soc. Am.* **82**, 1944–1956.
- Strickland, E. A., Viemeister, N. F., Fantini, D. A., and Garrison, M. A. (1989). "Within- versus cross-channel mechanisms in detection of envelope phase disparity," *J. Acoust. Soc. Am.* **86**, 2160–2166.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- van de Par, S., and Kohlrausch, A. (1998). "Analytical expressions for the envelope correlation of narrow-band stimuli used in CMR and BMLD research," *J. Acoust. Soc. Am.* **103**, 3605–3620.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.