Published in final edited form as:
J Acoust Soc Am. 2005 December ; 118(6): 3774–3782.

Detection of spectrally complex signals in comodulated maskers: Effect of temporal fringe

John H. Grose[a], **Joseph W. Hall III**, **Emily Buss**, and **Debora R. Hatch**
Department of Otolaryngology/Head & Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7070

Abstract

This study tested the hypothesis that masking release for a complex signal under conditions where signal energy is present in all frequency regions occupied by the masker is attributable to an across-frequency-channel comodulation masking release (CMR) process. The approach was to identify a signature CMR trait, and to then determine if that trait was associated with the detection advantage for complex signals. The selected trait was the decline of CMR in the presence of a random temporal fringe. In experiment 1, a masking release was observed for a 4-component harmonic signal presented in a comodulated masker, and this masking release was diminished by the random temporal fringe. A similar effect was observed in experiment 2 for a 4-component inharmonic signal. These results support the hypothesis that a CMR can be measured for a complex signal even when there is substantial spectral overlap between the signal and its comodulated masker. This finding has consequences for CMR models since it demonstrates that the presence of 'signal-free' cue bands is not a prerequisite for CMR, and that the presence of comodulation during the signal window is not sufficient to result in CMR.

I. INTRODUCTION

Comodulation masking release (CMR) refers to the detection advantage conferred on a masked signal when the masker exhibits a coherent amplitude fluctuation pattern across frequency channels relative to when the fluctuation pattern is independent across frequency channels. The operational phrase here is 'across frequency channel' since, as will be discussed in more detail below, the term CMR is not applicable to those situations where a masking release is observed for stimuli confined to a single auditory filter ('within frequency channel'). In most realizations of CMR, the comodulated masker occupies frequency regions that are relatively free of signal energy, and which therefore provide information about the waveform statistics that are unique to the masker. This configuration lends itself to models of CMR wherein frequency regions containing masker energy alone facilitate the identification of low-energy epochs of the masker where the instantaneous signal-to-noise ratio is relatively high (for review, see Hall et al., 1995; van de Par and Kohlrausch, 1998; Verhey et al., 2003). In physiological implementations of this approach, such signal enhancement effects can be realized by inhibitory neural circuits activated by frequency regions of the masker that are devoid of signal energy (Pressnitzer et al., 2001). Models such as these are not designed to deal with stimulus configurations where there is substantial spectral overlap between a complex signal and its comodulated masker.

The purpose of this study is to determine whether mechanisms underlying CMR play a role in auditory environments where signal energy exists in every frequency region where comodulated masking energy is present. This is an important question because, whereas most

[a] electronic mail: jhg@med.unc.edu.

laboratory demonstrations of CMR involve a pure-tone signal, ecologically significant signals are typically more complex. In a configuration where each component of a complex tonal signal is masked by a narrow band of noise centered at the component frequency, threshold for the complex signal will likely be lower if all masker bands are comodulated than if they are all independent (Hall et al., 1988; Bacon et al., 2002). However, this detection advantage does not necessarily represent an authentic CMR. Consider the study of Bacon et al. (2002) in which maskers were constructed of 100-Hz-wide bands of noise that were either unmodulated or were sinusoidally modulated at a low rate. For any single 100-Hz-wide band presented alone, threshold for a tone centered in that band was lower in the modulated case than in the unmodulated case – a detection advantage that was within-frequency-channel in nature. When the masker consisted of three widely spaced bands of noise and the signal was a 3-tone complex centered in those bands, a detection advantage was similarly observed for the complex signal in the comodulated masker relative to the unmodulated masker. However, because a signal threshold reduction had been demonstrated for each individual modulated masker band alone, the detection advantage measured for the complex signal masked by the comodulated noise-bands could not be attributed unambiguously to CMR. Rather, it could be argued that the signal component in each frequency channel was detected on a within-channel basis, and that threshold for the signal complex as a whole was determined by the spectral integration of the 'individually extracted' components. Herein lies the quandary: Given that threshold for a complex signal will likely be lower if all masker bands centered on the signal components are comodulated than if they are all independent, how can it be determined whether the detection advantage reflects a veridical across-frequency-channel CMR?

One approach to this issue is to identify a signature characteristic of CMR and to determine whether this trait can be observed in the detection advantage measured for the complex signals. To illustrate this approach, two examples will be described: the first deals with pure-tone signals (Dau et al., 2004); the second deals with complex-tone signals (Grose and Hall, 1992). Dau et al. (2004) measured pure-tone detection in maskers made up of 20-Hz-wide bands of noise that were either closely spaced in frequency to fall within a single frequency channel or were widely spaced to fall into relatively independent frequency channels. In each composite masker, the constituent noise-bands were either comodulated or random. For both the closely-spaced and widely-spaced composite maskers, detection threshold for a pure tone centered in the masker was lower in the comodulated masker than in the random masker. Based on the premise that CMR represents an across-frequency-channel process, it is likely that the detection advantage associated with the widely-spaced masker reflected an authentic CMR. However, by the same token, it is not likely that the detection advantage associated with the closely-spaced masker was an authentic CMR. The challenge faced by Dau et al. (2004) was distinguishing between the underlying mechanisms contributing to the observed detection advantages in the two cases. They noted that an authentic across-frequency-channel CMR was subject to auditory grouping effects, such that perceptual segregation of the flanking cue bands from the on-signal noise-band in the comodulated masker undermined CMR (cf. Grose and Hall, 1993). Based on this signature CMR trait, they reasoned that exposing both the closely-spaced and widely-spaced comodulated maskers to such manipulations should differentiate between the underlying mechanisms of masking release. Using asynchronies on the flanking cue bands to promote perceptual segregation, they demonstrated that the masking release was obliterated for the widely-spaced masker but not for the closely-spaced masker. Thus, they concluded that the signal detection advantage associated with the widely-spaced masker reflected an authentic CMR, whereas that for the closely-spaced masker did not.

The second example, using complex-tone signals, comes from the study of Grose and Hall (1992). They noted that the magnitude of CMR is sensitive to the number of comodulated noise-bands present, such that CMR increases asymptotically as the number of bands increases (Carlyon et al., 1989; Schooneveldt and Moore, 1989; Haggard et al., 1990; Hall et al., 1990;

Hatch et al., 1995). Taking this as a signature CMR trait, they demonstrated that the detection advantage measured for a 3-tone signal masked by three comodulated noise-bands was enhanced by the addition of (twelve) further comodulated noise-bands. They therefore concluded that the original detection advantage was likely due to CMR.

Both of these studies employed the logic that, (a) if a signature CMR trait can be identified, and (b) if this trait can be shown to be associated with the detection advantage observed for a particular signal, then (c) the observed detection advantage represents an authentic CMR. The same reasoning was used here to further investigate CMR under conditions where signal energy exists in every frequency region where comodulated masker energy is present. However, the manipulation of onset/offset synchrony used by Dau et al. (2004) was not well-suited to the present study because the concepts of 'on-signal band' and 'cue band' lose their distinction in the multi-component signal paradigm employed here. Similarly, the use of additional (signal-free) cue bands implemented by Grose and Hall (1992) was poorly suited to the present paradigm because this manipulation undermines the goal of assessing masking release in conditions where substantial spectral overlap exists between signal and masker energy. Instead, it was decided to pursue an effect first reported by Mendoza et al. (1998), but not formally published. In that study, it was noted that the magnitude of CMR for a pure-tone signal was highly sensitive to the characteristics of the noise that temporally bracketed the observation intervals containing the comodulated noise-bands. Specifically, when only the gated comodulated masker was present, a substantial CMR was observed; however, when the gated comodulated masker was embedded in a temporal fringe that consisted of narrow bands of noise that were in all respects similar to the masker except that their envelopes were random across frequency (not comodulated), CMR declined markedly. The sensitivity of CMR to the random noise-bands temporally bracketing the comodulated maskers was concluded to be a signature trait of CMR processing. In the context of this investigation, the attraction of this CMR trait is that it does not require the inclusion of additional masker energy at other frequencies, and does not require gating asynchronies across the comodulated noise-bands.

In summary, the purpose of this investigation was to determine whether mechanisms underlying CMR are applicable to situations where signal energy exists in every frequency region where comodulated masking energy is present. The detectability of individual signal components was equated, and the effect of a temporal fringe comprised of random noise-bands was taken as an indication of whether any detection advantage observed could be attributed to a valid CMR process. The hypothesis was that a masking release would be observed for complex signals in a comodulated masker, and that this would reflect an authentic CMR process as demonstrated by the sensitivity of the masking release to a random temporal fringe. Two experiments were undertaken. The first experiment incorporated harmonic maskers and referenced CMR to detection thresholds in a single narrow band of noise. The second experiment incorporated inharmonic maskers and referenced CMR to detection thresholds in random maskers.

II. EXPERIMENT 1. EFFECT OF A RANDOM TEMPORAL FRINGE ON COMPLEX SIGNAL DETECTION IN COMODULATED HARMONIC MASKERS

The purpose of experiment 1 was to demonstrate that a masking release could be observed for a complex signal masked by a set of comodulated narrow bands of noise, where each band of noise was centered at a signal component and where the detectability of each signal component was equated. The hypothesis was that this masking release is due to CMR mechanisms as evidenced by the reduction in the magnitude of masking release due to the presence of a random temporal fringe. A four-component complex signal was employed and, hence, the on-component masker consisted of four narrow bands of noise. As a supplementary test of CMR, a 10-band masker was also included since, as noted earlier, the magnitude of CMR is known to increase with the number of comodulated noise-bands.

A. Method

1. Subjects—Eight listeners ranging in age from 19 to 46 years (mean = 28.3 years) participated. All had normal hearing, with audiometric thresholds below 20 dB HL across the octave frequencies 250–8000 Hz (ANSI, 1996). All listeners practiced the task until thresholds appeared stable; this typically took less than 1 hour.

2. Stimuli—The signals were pure tones having frequencies of 1600, 2000, 2400, or 2800 Hz, as well as a complex tone consisting of all four frequencies. Each signal was 400 ms in duration, including 50-ms cosine-squared rise/fall ramps. Three maskers were constructed. The first masker consisted of a single 20-Hz-wide band of noise centered at 2000 Hz. The second masker consisted of four 20-Hz-wide, comodulated bands of noise centered at 1600, 2000, 2400 and 2800 Hz. Complementing this masker was a set of four 20-Hz-wide, independent bands of noise centered at the same frequencies. The third masker consisted of 10 comodulated bands of noise, each 20-Hz-wide, centered at the first ten harmonics of 400 Hz (400–4000 Hz). Again, complementing this masker was a set of 10 independent 20-Hz-wide bands of noise centered at the same frequencies. The frequency region of the signal and the harmonic structure of the stimuli followed from the earlier work of Mendoza et al. (1998). All bands of noise were generated by an inverse FFT technique using a long buffer size that resulted in a temporal waveform approximately 6.6 s in duration that wrapped seamlessly because all noise components consisted of an integer number of cycles. For any particular condition, the appropriate comodulated masker (4-band or 10-band) was played continuously out of one channel of a DAC (TDT PD1), while its complementary set of random bands was played continuously out of a separate DAC channel. The noise-band sets were gated on and off as needed using electronic switches which imposed 50-ms cosine-squared rise/fall ramps (TDT SW2). All stimuli were generated at a sampling rate of 20 kHz and were antialias filtered at 8 kHz (Kemo VBF8). The maskers were presented at a spectrum level of 52 dB/Hz SPL (65 dB SPL per noise-band) through the left phone of a Sennheiser HD580 headset.

3. Procedure—Signal detection threshold was measured using a three-alternative, forced-choice (3AFC) procedure incorporating a three-down, one-up stepping rule that estimated the 79.4% correct point on the psychometric function. The observation intervals were marked with LEDs on a hand-held response box and the listeners entered their interval selection by means of a corresponding button on the same box. The initial step-size of signal level adjustment was 4 dB and, following each of the first two reversals in signal level direction, the step-size was halved. A threshold estimation track was terminated after eight reversals in level direction, and threshold was computed as the mean of the final six reversal levels. For each condition, four estimates of threshold were obtained, and the final threshold value was taken as the mean of all four estimates.

There were three masking conditions that formed the core of the experiment: (1) a baseline condition against which masking release was referenced; (2) a gated comodulated condition where each of the four components of the complex signal was masked by an on-component comodulated noise-band; and (3) a temporal fringe condition where the gated comodulated masker of (2) was bracketed by its complementary set of four random noise-bands. In addition to these three main conditions, other conditions were implemented where the number of noise-bands (four or ten) and the nature of the temporal fringe (random or comodulated) were varied to test aspects of the hypothesis. The rationale for these manipulations will be noted at the relevant points in the discussion below.

The choice of baseline for this experiment was the threshold of the 2000-Hz pure-tone signal masked by the single 20-Hz wide band of noise centered at 2000 Hz (in this context, termed the on-signal band [OSB]). In CMR experiments where the signal consists of a pure tone, this baseline is an accepted reference against which to measure masking release, and the derived release is sometimes termed the Reference-Comodulated (R-C) measure of CMR (Schooneveldt and Moore, 1987). The use of this baseline here rests partly on the assumption that the threshold of the 2000-Hz signal masked by its OSB is representative of the thresholds that would have been measured for the other pure-tone signals masked by their respective on-signal bands. This assumption is valid since it has been shown that pure-tone thresholds in narrow on-signal masker bands do not vary significantly with frequency (Bos and de Boer, 1966). However, the validity of using this baseline in an experiment where the main signal of interest is a four-component complex requires further support because of the expected reduction in signal threshold due to spectral integration. As expanded on in the footnote, the reduction in signal threshold due to spectral integration under the conditions of this experiment is countered by an increase in threshold due to across-channel masking.[1] Thus the OSB baseline provides a viable reference. The OSB masker was gated on for 500 ms in each observation interval of the 3AFC trial, and the 2000-Hz signal was gated on in the temporal center of one of the observation intervals at random. The inter-stimulus interval within each 3AFC trial was 500 ms.

The second main condition consisted of the 4-band comodulated masker, which was gated on with the same timing as in the baseline condition. Detection thresholds were measured first for the four individual pure-tone signals. Once threshold levels for these signals had been determined for a given listener, a complex-tone signal was constructed that was tailored to that listener’s thresholds. The levels of the individual components within the 4-tone complex were adjusted to provide equal detectability. This was accomplished by attenuating components relative to the component with the highest masked pure-tone threshold by an amount that corresponded to the respective pure-tone threshold difference between them. For example, if the highest masked threshold was 63 dB SPL for the 1600-Hz tone, whereas the threshold for the 2000-Hz tone was 61 dB SPL, then within the 4-tone complex the 2000-Hz component would be fixed at 2-dB down from the 1600-Hz component. Threshold for the complex signal was then referenced to the level of the most intense component. Note that in the complex signal, component starting phase was randomized across components.

In the third main condition, a temporal fringe was added to the gated comodulated masker. The temporal fringe consisted of a continuous presentation of the four random noise-bands except during the observation intervals of the 3AFC trial where the comodulated masker was gated on. The transition between the surrounding random noise and the gated comodulated masker consisted of overlapping 50-ms rise/fall ramps; i.e., the surrounding noise was gated off as the

---

[1] With regard to the complex signal, a valid objection to the R-C method of deriving CMR in experiment 1 is that it potentially overestimates the CMR magnitude by ignoring any effect of spectral integration due to the presence of multiple signal components. That is, if the reference for the complex signal had been its threshold in a masker composed of independent bands of noise (one band centered on each signal component), rather than a single tone masked by a single band of noise, then the reference threshold might have been lower due to spectral integration. A lower reference threshold would have resulted in a smaller derived measure of CMR. Whereas this objection is sound, it fails to take into account any increase in the reference threshold for the pure-tone signals due to across-channel masking associated with the multiple masker bands (Hall et al., 1984; McFadden, 1986; Cohen and Schubert, 1987; Schooneveldt and Moore, 1987; Eddins and Wright, 1994). That is, failure to incorporate potential effects of across-channel masking in defining the reference threshold in experiment 1 would have resulted in an underestimation of the CMR magnitude. To illustrate this, additional data were collected from two listeners where the reference masker consisted of the four independent bands of noise centered at the signal component frequencies. Thresholds for the 2000-Hz pure-tone signal measured in the masker were, on average, 2.1 dB higher for the two listeners than their thresholds for the 2000-Hz signal masked by its on-signal band alone (the reference used in experiment 1). Spectral integration was indeed observed for the complex signal in the random reference masker (on average, 3.1 dB for the two listeners), but the threshold reduction due to spectral integration was largely offset by the threshold elevation due to across-channel masking. Thus, the actual reference used in experiment 1 is a viable baseline against which to measure CMR, even for the complex signal. In any case, the results of experiment 1 are substantiated by experiment 2 which did use the multiple random noise bands as the reference for masking release.

comodulated masker was gated on at the beginning of an observation interval, and the surrounding noise was gated on as the comodulated masker was gated off at the termination of the observation interval. This is illustrated in Fig. 1. The overlapping switching yielded perceptually seamless transitions between the two noise sources. For the temporal fringe condition, detection thresholds were measured first for the four separate pure-tone signals followed by threshold for an individually tailored complex-tone signal. (Note that an individually tailored complex signal was made separately for the gated comodulated condition and the temporal fringe condition.)

In addition to these three main conditions, data were also collected for other conditions where either the number of noise-bands (4 or 10) or the characteristics of the temporal fringe (random or comodulated) were varied. In all these conditions, the gating patterns were the same as described above. Again, for any masking condition that involved the complex signal, the relative levels of the four components were individually tailored for each listener for that masker; i.e., for each masker the pure-tone thresholds were measured first and the complex signal then individually constructed. For each listener, the conditions were blocked, but the order of conditions was random across listeners.

B. Results and Discussion

The basic pattern of results was similar across the eight listeners and so the group means are representative of all listeners' performance. The mean results for the three main masking conditions are shown in Fig. 2. To highlight the effect first reported by Mendoza et al. (1998), consider first the data pattern associated with the 2000-Hz signal. The average threshold in the baseline OSB masker was about 67 dB SPL (triangle). Detection threshold decreased to about 61 dB SPL in the 4-band comodulated masker (circle), yielding a CMR of about 6 dB. When the comodulated masker was presented within its complementary random temporal fringe, thresholds rose to about 65 dB SPL (square), reducing CMR to 2 dB. The purpose of experiment 1 was to determine whether this pattern was real and present for all signals, particularly the complex signal (highlighted by unfilled symbols in Fig. 2). The first step in the analysis was to determine whether a masking release was associated with signal detection in the 4-band comodulated masker. With reference to Fig. 2, this analysis sought to determine whether the thresholds represented by the filled circles were significantly lower than the reference threshold shown by the horizontal bar (the threshold for the 2000-Hz signal masked by its OSB, with the shaded region indicating ±1 standard deviation). A repeated-measures analysis of variance (ANOVA) was undertaken that included thresholds for the five signal types tested with this masker, as well as the reference threshold. The analysis revealed a significant effect of condition ($F_{5,35} = 7.943$; $p < 0.0001$), and post-hoc simple contrasts indicated that detection thresholds for all signal types in the 4-band comodulated masker differed significantly from the OSB reference, except for the 1600-Hz signal. Thus, a masking release was observed in the 4-band comodulated masker for all signals, including the complex signal, with the exception of the 1600-Hz signal. Failure to observe a significant CMR for the 1600-Hz signal is likely due to its spectral position on the lower edge of the stimulus complex. Hall et al. (1988) have shown that, for a set of comodulated narrow-band maskers, CMR is minimal for a signal centered in the lowest frequency band. For the remaining signals, the magnitude of masking release ranged from 3.8 dB for the 2800-Hz signal to 6.3 dB for the 2000-Hz signal.

The key experimental question was whether this masking release diminished in the presence of the random temporal fringe, particularly for the complex signal. To assess this, a repeated-measures ANOVA was undertaken that compared the thresholds in the 4-band gated comodulated masker to those obtained in the condition where the 4-band gated comodulated masker was embedded in a temporal fringe consisting of four random noise-bands. With

NIH-PA Author Manuscript
NIH-PA Author Manuscript

reference to Fig. 2, this analysis sought to determine whether the thresholds represented by the squares were significantly higher than the thresholds represented by the circles. Data for the 1600-Hz pure-tone signal were excluded from this analysis since no masking release was observed for this signal. The analysis incorporated two within-subjects factors: masker configuration (comodulated vs. comodulated + fringe) and signal type. The analysis indicated a significant reduction in masking release in the presence of the random temporal fringe ($F_{1,7} = 11.88$; $p = 0.011$) with no effect of signal type and no interaction between these factors. Thus, the effect noted by Mendoza et al. (1998) held for all signal types for which a masking release had been observed, including the complex signal.

The motivation of this study was to determine whether the detection advantage for the complex signal in the comodulated masker represented an authentic CMR. The experimental approach was to demonstrate the sensitivity of the masking release to the presence of a random temporal fringe, a manipulation known to be detrimental to CMR. The results therefore support the interpretation that the masking release measured for the complex signal is the product of CMR mechanisms. In order to further substantiate this interpretation, two supplementary tests were undertaken, each testing another specific feature of CMR: (1) the effect of the number of cue bands; and (2) the effect of continuous vs. gated presentation of the comodulated noise-bands.

As noted earlier, it is known that the magnitude of CMR is sensitive to the number of comodulated noise-bands (Carlyon et al., 1989; Schooneveldt and Moore, 1989; Haggard et al., 1990; Hall et al., 1990; Hatch et al., 1995). If the threshold reduction for the complex signal in the 4-band comodulated masker is due to CMR, then threshold should be reduced still further upon an increase in the number of comodulated noise-bands. This trait was tested for here by measuring the change in threshold for the complex signal in the 4-band comodulated masker when six additional cue bands were added. The average threshold in the 4-band masker was 62.7 dB SPL and this was reduced to a mean of 58.4 dB SPL in the 10-band comodulated masker. A paired t-test indicated that this reduction was significant ($t_7 = 2.579$; $p = 0.037$).

In terms of gating effects, it has been shown that for a 2000-Hz signal masked by a relatively few number of comodulated noise-bands (3–5), detection thresholds improve by about 6 dB when the masker is presented continuously rather than gated synchronously with the signal (Hatch et al., 1995). In the present context, this is equivalent to presenting the gated 4-band comodulated masker in a temporal fringe consisting of the same four comodulated noise-bands. As noted above, the average threshold for the complex signal in the 4-band gated masker was 62.7 dB SPL and this decreased to 58.5 dB SPL in the continuous masker. A paired t-test showed this reduction to be significant ($t_7 = 4.174$; $p = 0.004$). These two supplementary tests therefore add further support to the notion that the masking release observed for the complex signal in the 4-band gated comodulated masker is a product of CMR mechanisms.

In summary, the results of this experiment support the notion that CMR can play a role in auditory environments where signal energy exists in every frequency region where comodulated masking energy is present. The sensitivity of CMR to the characteristics of the noise bracketing the comodulated observation intervals corroborates existing findings that show relatively long-term effects of stimulus coherence. For example, introducing an onset asynchrony between the OSB masker and the comodulated flanking bands reduces CMR despite the fact that all maskers are present and comodulated during the observation intervals (Grose and Hall, 1993). Such long-term effects of stimulus coherence have been demonstrated over a wide range of stimuli and phenomena (Hall and Grose, 1990; Grose and Hall, 1993; Oxenham and Dau, 2001; Dau et al., 2004). In terms of CMR, they are likely to represent a general characteristic of the phenomenon.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

As noted earlier, there are traditionally two methods of deriving CMR. The method used in this experiment, the R-C derivation, uses the threshold of a pure-tone signal masked by its OSB as the reference. The second method of deriving CMR, known as the U-C derivation, uses as the baseline the signal threshold in the full complement of random noise-bands making up the masker (Schooneveldt and Moore, 1987). Because the use of a complex signal makes the optimum choice for baseline less straightforward, as discussed in the footnote, a second experiment was undertaken that employed the U-C method of deriving CMR.

III. EXPERIMENT 2. EFFECT OF A RANDOM TEMPORAL FRINGE ON COMPLEX SIGNAL DETECTION IN COMODULATED INHARMONIC MASKERS

The purpose of experiment 2 was again to demonstrate that a masking release could be observed for a complex signal masked by a set of comodulated narrow bands of noise under conditions where there was one noise-band per signal component and where the detectability of each signal component was equated. The hypothesis was that this masking release is due to CMR mechanisms as evidenced by the reduction in the magnitude of masking release due to the presence of a random temporal fringe. A four-component complex signal was employed and, hence, the on-component masker consisted of four narrow bands of noise. In addition to employing the U-C method of deriving CMR, experiment 2 made use of inharmonic spacings between signal components. Because harmonicity is a strong auditory grouping cue (e.g., Bregman et al., 1990; Darwin et al., 1991), the motivation was to undermine the contribution of this cue to the perceptual organization of the comodulated noise-bands; i.e. it was desired that the primary grouping cue for the noise-bands would be their temporal envelope coherence.

A. Method

1. Subjects—Seven listeners ranging in age from 19 to 47 years (mean = 30.4 years) participated in the second experiment. All had normal hearing, with audiometric thresholds below 20 dB HL across the octave frequencies 250–8000 Hz (ANSI, 1996). All listeners practiced the task using a pure-tone signal until thresholds appeared stable; this typically took less than 1 hour.

2. Stimuli—The signals were pure tones having frequencies of 804, 1200, 1747, or 2503 Hz, as well as a complex tone consisting of all four frequencies. These frequencies were selected as being the center frequencies of equivalent rectangular bandwidths [ERBs] (Moore and Glasberg, 1987) that formed a series in which two non-overlapping ERBs intervened between each selected ERB. The frequencies thus activated relatively discrete regions of the basilar membrane. As in experiment 1, signals were 400 ms in duration, including 50-ms cosine-squared rise/fall ramps. The masker consisted of four 20-Hz-wide narrow bands of noise centered at these frequencies; the bands were either comodulated or random with respect to each other. Each four-band masker (comodulated or random) was generated by the same inverse FFT technique described earlier. The comodulated masker was output continuously through one channel of the DAC (TDT PD1) whereas the random masker was output through the second channel. The noise-band sets were gated on and off as needed using electronic switches which imposed 50-ms cosine-squared rise/fall ramps (TDT SW2). Whereas the maskers in experiment 1 were presented at a level of 65 dB SPL per noise-band, here the overall level of the 4-band masker was 65 dB SPL (c. 46 dB/Hz).

3. Procedure—The procedure for measuring signal detection threshold was the same as that in experiment 1 except that five estimates of threshold were obtained for each condition. Final threshold value was taken as the mean of all five estimates. Threshold for each signal type was measured in three masking conditions: (1) the gated random masker [baseline]; (2) the gated comodulated masker; and (3) the gated comodulated masker bracketed by a temporal fringe consisting of the random noise bands. For each of these maskers, thresholds for the pure-tone signals were measured first and then the complex signal was individually tailored for each listener based on the pure-tone thresholds. For the masking condition with the temporal fringe, the transition between the surrounding random noise and the gated comodulated masker consisted of overlapping 50-ms rise/fall ramps as in experiment 1. For each listener, the conditions were blocked, but the order of conditions was random across listeners, with the proviso that for any particular masker condition the pure-tone thresholds were measured prior to the complex-tone thresholds.

B. Results and Discussion

The basic pattern of results was similar across the seven listeners and so the group means are representative of all listeners' performance. The mean results for each signal and masking condition are shown in Fig. 3. Thresholds for the gated, random masker (triangles) were higher than thresholds for the gated, comodulated masker (circles), although the magnitude of the difference appeared to depend on signal type. (Again, results for the complex signal are highlighted as unfilled symbols.) To test this observation, a repeated-measures ANOVA was undertaken on the gated thresholds using two within-subjects factors: signal type (5 levels) and masker type (2 levels). The analysis indicated that the main effect of signal type was significant ($F_{4,24} = 2.965$; $p = 0.04$), as was the effect of masker type ($F_{1,6} = 48.619$; $p < 0.001$); however, the interaction between signal type and masker type was also significant ($F_{4,24} = 4.45$; $p = 0.008$). To interpret the interaction between signal and masker types, pair-wise comparisons between the signal thresholds obtained in the two types of masker were undertaken for each signal type using paired t-tests with Bonferroni correction. The results of the pair-wise analysis indicated that thresholds were reliably lower in the comodulated masker for all signal types except for the 804-Hz pure tone signal. As with the 1600-Hz signal of experiment 1, the failure to observe a masking release for the 804-Hz signal here is probably due to its spectral position on the low-frequency edge of the stimulus configuration. The magnitude of the masking release for the remaining signals ranged from 3.7 dB for the 2503-Hz signal to 6.3 dB for the 1747-Hz signal. Note that a significant threshold reduction of about 4.4 dB was apparent also for the complex-tone signal.

The key issue was the effect of a random temporal fringe on the masking release observed for the gated comodulated conditions. The results of this condition are shown as squares in Fig. 3. To determine the effect of this condition, a repeated-measures ANOVA was undertaken that incorporated two within-subjects factors: masker configuration (comodulated vs. comodulated + fringe) and signal type. Only the four signal types that were associated with a masking release were included in the analysis; i.e., the 804-Hz data were excluded. The analysis revealed that the effect of masker configuration was significant ($F_{1,6} = 7.498$; $p = 0.034$). There was no effect of signal type, nor was the interaction between signal type and masker configuration significant. The results of this analysis therefore indicate that the presence of a random temporal fringe significantly elevated signal threshold (reduced CMR) in those configurations where the gated comodulated noise-bands initially led to a masking release. Because this occurred also for the masking release observed for the complex signal, it can be concluded that this masking release reflects the contributions of CMR mechanisms.

Whereas the pattern of results supports an interpretation in terms of the role of CMR processing, it does not rule out accounts framed in terms of spectral integration. However, such accounts would require that spectral integration be sensitive to the temporal fringe of the masker. Evidence against this can be gleaned from two supplementary conditions tested on a subset of listeners in experiment 2. These two conditions examined the effect of a temporal fringe on a masker configuration that is not conducive to CMR, i.e., a gated random masker. Three of the listeners were tested with a comodulated temporal fringe bracketing the gated random masker, and four of the listeners were tested with a random temporal fringe bracketing the gated random masker (equivalent to a continuous random masker). Neither of these conditions resulted in a significant change in threshold relative to that measured in the gated random masker alone. This supports the notion that spectral integration is not sensitive to temporal fringe. In turn, this interpretation reinforces the contention that the effect of the random temporal fringe on thresholds for the complex signal in comodulated noise indicates the involvement of CMR mechanisms.

IV. GENERAL DISCUSSION

Confirmation that CMR processing can take place even when signal and masker spectra substantially overlap is informative as to the nature of CMR. As noted by Hall et al. (1988), the fact that the presence of ‘signal-free’ cue bands is not a prerequisite for the occurrence of CMR does not appear at first sight to be compatible with dip-listening strategies. A number of studies have provided evidence that the underpinning of CMR is the ability to preferentially monitor the epochs of low masker energy (i.e., ‘dips’) where the signal-to-noise ratio is relatively optimal (Buus, 1985; Grose and Hall, 1989; Moore et al., 1990; Hall and Grose, 1991; Hicks and Bacon, 1995; Buus et al., 1996). Phenomenologically, this benefit is described in terms of a process wherein the comodulated flanking bands cue the auditory system as to when to ‘listen’ for the signal. Such a description implies the necessity of ‘signal-free’ cue bands. However, signal enhancement, or increased weighting of the stimulus energy during the masker dips, will result from any process in which the envelope information associated with comodulated cue bands is used to account for the envelope information originating from the masker in the signal channel. Physiologically, Pressnitzer et al. (2001) have modeled this as a neural circuit in which the output of a cell with a wide-band receptive field responding to all the cue bands leads to a fast-acting inhibition of a cell with a narrow-band receptive field that responds to only the signal frequency. If a compressive function is added to the circuit to institute some degree of equalization between the inhibitory and excitatory inputs to the narrow-band cell, then this circuit provides a simple equalization/cancellation mechanism that can result in CMR-like behavior. A further feature of an equalization/cancellation mechanism, of which the Pressnitzer model is a particular example, is that even if a signal component exists in every frequency channel conveying comodulated masker information, such a mechanism could potentially have a higher output when the signal was present than when just the comodulated masker alone was present. That is, a higher residue would remain at the output of the cancellation stage when a signal was present than when a signal was absent because, with the random starting phases of the individual signal components, the envelopes of the signal + masker waveforms in the various frequency channels would not be identical. This higher output may in turn constitute a detection cue. Such a mechanism would be of no benefit for a random masker since it would have a significant output whether or not the signal was present. In this sense, the mechanism is informative only when the auditory system is anticipating coherent fluctuations across frequency. Perhaps this is why the presence of an independent temporal fringe reduces the benefit of comodulation, as observed here and in the Mendoza et al. (1998) study: the on-going random noise puts the system in a state where comodulation is not expected, and therefore potential cueing mechanisms (perhaps based on grouping by common modulation) are not activated. In line with this, Grose and Hall (1993) have shown that the presence of comodulation during the signal epoch is, in itself, insufficient to result in a CMR.

Although this study was not motivated to examine the characteristics of spectral integration under conditions of comodulation, the results are relevant to the findings of Bacon et al. (2002). That study specifically examined spectral integration for complex signals presented in narrow (100-Hz wide) bands of frozen noise that were either unmodulated, or were sinusoidally modulated at 8 Hz. They found that for unmodulated noise-bands, or for widely-spaced comodulated noise-bands, spectral integration reflected the expected $10\log\sqrt{N}$ rule, where N is the number of equally-detectable signal components. However, for noise-band spacings that were an octave or less, ‘super-integration’ was observed; i.e., amounts of spectral integration in excess of that predicted by the $10\log\sqrt{N}$ rule. They conjectured that, under these conditions, a CMR mechanism may act in tandem with a spectral integration process, although it is not clear how these additive effects would come about.

In order to compare the pattern of spectral integration observed in this study with that of Bacon et al. (2002), measures of spectral integration were derived for the multi-component signals used here. Because threshold level for the multi-component signal was referenced to the component with the highest pure-tone threshold, the derivation of spectral integration consists of a simple subtraction of the multi-component signal threshold from the highest pure-tone signal threshold for that masker set. Note that a $10\log\sqrt{N}$ rule would predict a 3-dB improvement in threshold between a pure-tone signal and a 4-component signal. From the data of experiment 1, spectral integration was computed for the two main conditions (4-band gated comodulated masker and the 4-band gated comodulated masker presented in a random temporal fringe), as well as the two supplementary conditions (10-band gated comodulated masker and the 4-band continuous comodulated masker). From the data of experiment 2, spectral integration was computed for the gated random masker, the gated comodulated masker, and the gated comodulated masker with the random temporal fringe. The results are shown in Fig. 4, and it can be seen that, on average, a 3-dB spectral integration was observed for three of the conditions, with less than this amount observed for the other conditions. However, individual variability was pronounced. Nevertheless, the present results do not appear to be in line with Bacon et al. (2002) in that no ‘super-integration’ was seen across all listeners for the comodulated bands that were separated by less than an octave from each other. It is possible that the marked differences in masker stimuli across the two studies may underlie this disparity. Here, the comodulation was dictated by the common inherent fluctuations across frequency in the envelopes of the 20-Hz-wide bands of noise, whereas in the Bacon et al. (2002) study the comodulation was dictated by the imposed 8-Hz sinusoidal modulation of the 100-Hz-wide bands of frozen noise. Further work is required to clarify this issue.

V. CONCLUSIONS

The purpose of this investigation was to test the hypothesis that a masking release would be observed for a complex signal in a comodulated masker, and that this masking release is attributable to a CMR process as demonstrated by the sensitivity of the masking release to a random temporal fringe. In both experiments 1 and 2, a masking release was observed for a complex signal, and this release was shown to be sensitive to the presence of a random temporal fringe. This finding provides support for the hypothesis that the masking release for the complex signal is associated with CMR processing. This conclusion has consequences for models of CMR since (a) it is demonstrated that the presence of ‘signal-free’ cue bands is not a prerequisite for the occurrence of CMR, and (b) the presence of comodulation during the signal window is not sufficient to result in a CMR.

Acknowledgements

This work was supported by NIH NIDCD (#R01-DC01507).

References

ANSI. American National Standards Specification for Audiometers (ANSI S3-1996). American National Standards Institute; 1996.

Bacon SR, Grimault N, Lee J. Spectral integration in bands of modulated or unmodulated noise. J Acoust Soc Am 2002;112:219–26. [PubMed: 12141347]

Bos CE, de Boer E. Masking and discrimination. J Acoust Soc Am 1966;39:708–715.

Bregman AS, Levitan R, Liao C. Fusion of auditory components: Effects of the frequency of amplitude modulation. Percept Psychophys 1990;47:68–73. [PubMed: 2300426]

Buus S. Release from masking caused by envelope fluctuations. J Acoust Soc Am 1985;78:1958–1965. [PubMed: 4078172]

Buus S, Zhang L, Florentine M. Stimulus-driven, time-varying weights for Comodulation Masking Release. J Acoust Soc Am 1996;99:2288–2297. [PubMed: 8730075]

Carlyon RP, Buus S, Florentine M. Comodulation masking release for three types of modulators as a function of modulation rate. Hear Res 1989;42:37–46. [PubMed: 2584158]

Cohen MF, Schubert ED. Influence of place synchrony on detection of a sinusoid. J Acoust Soc Am 1987;81:452–458. [PubMed: 3558962]

Darwin, CJ.; Buffa, A.; Williams, D.; Ciocca, V. Pitch of dichotic complex tones with a mistuned frequency component. In: Cazals, Y.; Demany, L.; Horner, K., editors. Auditory Physiology and Perception. Pergamon; Oxford: 1991.

Dau, T.; Ewert, SD.; Oxenham, AJ. Effects of concurrent and sequential streaming in comodulation masking release. In: Pressnitzer, D.; de Cheveigne, A.; McAdams, S.; Collet, L., editors. Auditory Signal Processing: Physiology, Psychoacoustics, and Models. Springer Verlag; 2004.

Eddins DA, Wright BA. Comodulation masking release for single and multiple rates of envelope fluctuation. J Acoust Soc Am 1994;96:3432–3442. [PubMed: 7814760]

Grose JH, Hall JW. Comodulation masking release using SAM tonal complex maskers: Effects of modulation depth and signal position. J Acoust Soc Am 1989;85:1276–1284. [PubMed: 2708670]

Grose JH, Hall JW. Comodulation masking release for speech stimuli. J Acoust Soc Am 1992;91:1042–1050. [PubMed: 1556306]

Grose JH, Hall JW III. Comodulation masking release: Is comodulation sufficient? J Acoust Soc Am 1993;93:2896–2902. [PubMed: 8315153]

Haggard MP, Hall JW, Grose JH. Comodulation masking release as a function of bandwidth and test frequency. J Acoust Soc Am 1990;88:113–118. [PubMed: 2380439]

Hall JW, Grose JH. Consequences of onset asynchrony for modulation detection interference. J Acoust Soc Am 1990;88:S145.

Hall JW, Grose JH. Relative contributions of envelope maxima and minima to comodulation masking release. Q J Exp Psychol 1991;43A:349–372.

Hall JW, Grose JH, Haggard MP. Comodulation masking release for multi-component signals. J Acoust Soc Am 1988;83:677–686. [PubMed: 3351126]

Hall JW, Grose JH, Haggard MP. Effects of flanking band proximity, number, and modulation pattern on comodulation masking release. J Acoust Soc Am 1990;87:269–283. [PubMed: 2299038]

Hall, JW.; Grose, JH.; Mendoza, L. Across-channel processes in masking. In: Moore, BCJ., editor. Handbook of Perception and Cognition Vol VI: Hearing. Academic Press; San Diego: 1995.

Hall JW, Haggard MP, Fernandes MA. Detection in noise by spectro-temporal pattern analysis. J Acoust Soc Am 1984;76:50–56. [PubMed: 6747111]

Hatch DR, Arne' BC, Hall JW. Comodulation masking release (CMR): Effects of gating as a function of number of flanking bands and masker bandwidth. J Acoust Soc Am 1995;97:3768–3774. [PubMed: 7790654]

Hicks ML, Bacon SP. Some factors influencing comodulation masking release and across-channel masking. J Acoust Soc Am 1995;98:2504–14. [PubMed: 7593933]

McFadden D. Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band. J Acoust Soc Am 1986;80:1658–1667. [PubMed: 3794071]

Mendoza L, Hall JW III, Grose JH. Comodulation masking release: the effect of the characteristics of noisebands presented before and after a signal. J Acoust Soc Am 1998;103:2843.

Moore BCJ, Glasberg BR. Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns. Hear Res 1987;28:209–225. [PubMed: 3654390]

Moore BCJ, Glasberg BR, Schooneveldt GP. Across-channel masking and comodulation masking release. J Acoust Soc Am 1990;87:1683–1694. [PubMed: 2341672]

Oxenham AJ, Dau T. Modulation detection interference: effects of concurrent and sequential streaming. J Acoust Soc Am 2001;110:402–8. [PubMed: 11508965]

Pressnitzer D, Meddis R, Delahaye R, Winter IM. Physiological correlates of comodulation masking release in the mammalian ventral cochlear nucleus. J Neurosci 2001;21:6377–6386. [PubMed: 11487661]

Schooneveldt GP, Moore BCJ. Comodulation masking release (CMR): effects of signal frequency, flanking band frequency, masker bandwidth, flanking band level, and monotic vs dichotic presentation flanking bands. J Acoust Soc Am 1987;82:1944–1956. [PubMed: 3429732]

Schooneveldt GP, Moore BCJ. Comodulation masking release (CMR) for various monaural and binaural combinations of the signal, on-frequency, and flanking bands. J Acoust Soc Am 1989;85:262–272. [PubMed: 2921408]

van de Par S, Kohlrausch A. Comparison of monaural (CMR) and binaural (BMLD) masking release. J Acoust Soc Am 1998;103:1573–9. [PubMed: 9514021]

Verhey JL, Pressnitzer D, Winter IM. The psychophysics and physiology of comodulation masking release. Exp Brain Res 2003;153:405–17. [PubMed: 13680049]

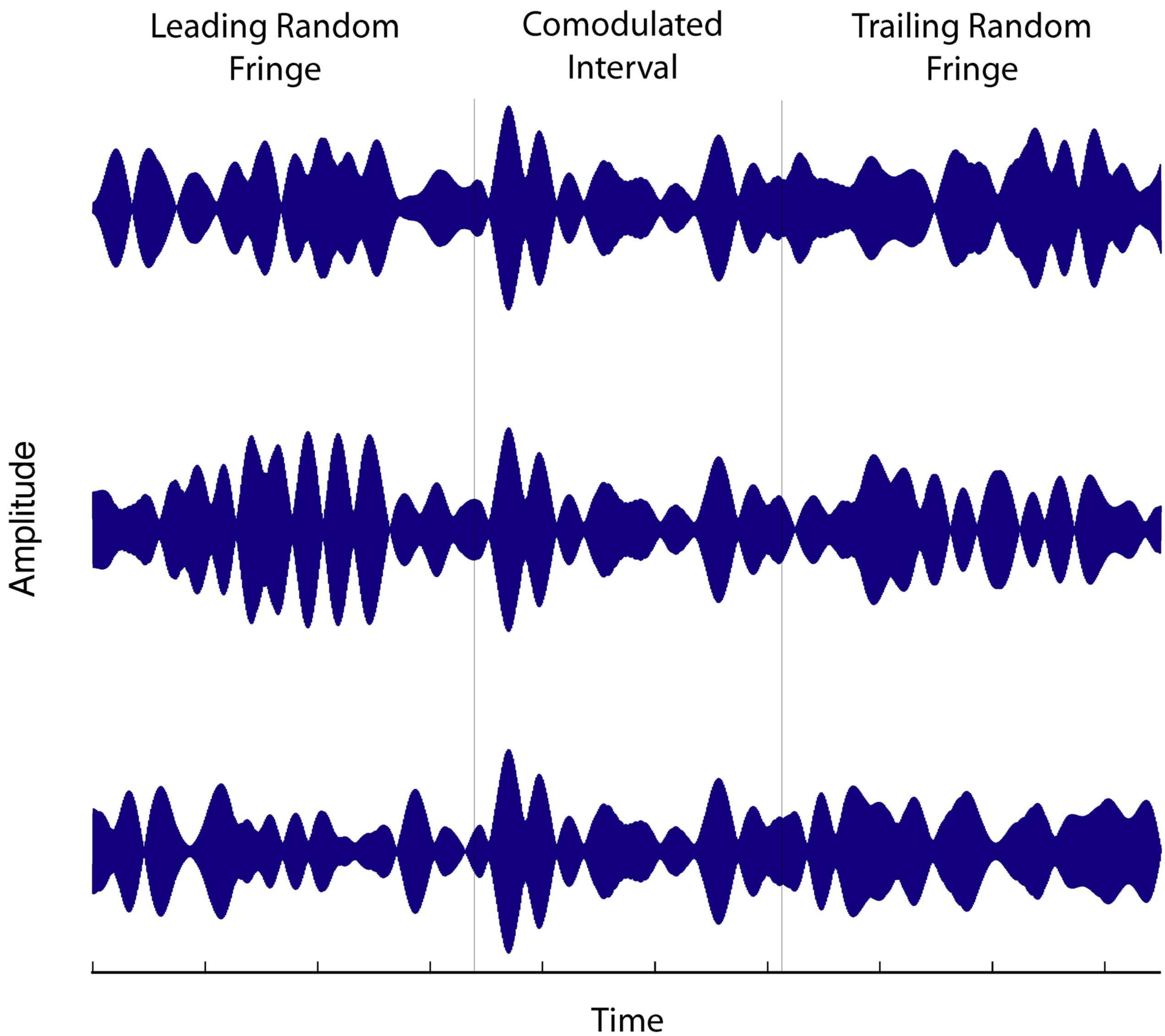

Figure 1.
Schematic of a comodulated observation interval with a random temporal fringe. Three noise-bands are depicted whose envelopes are uncorrelated before and after the observation interval, but whose envelopes are comodulated during the observation interval.

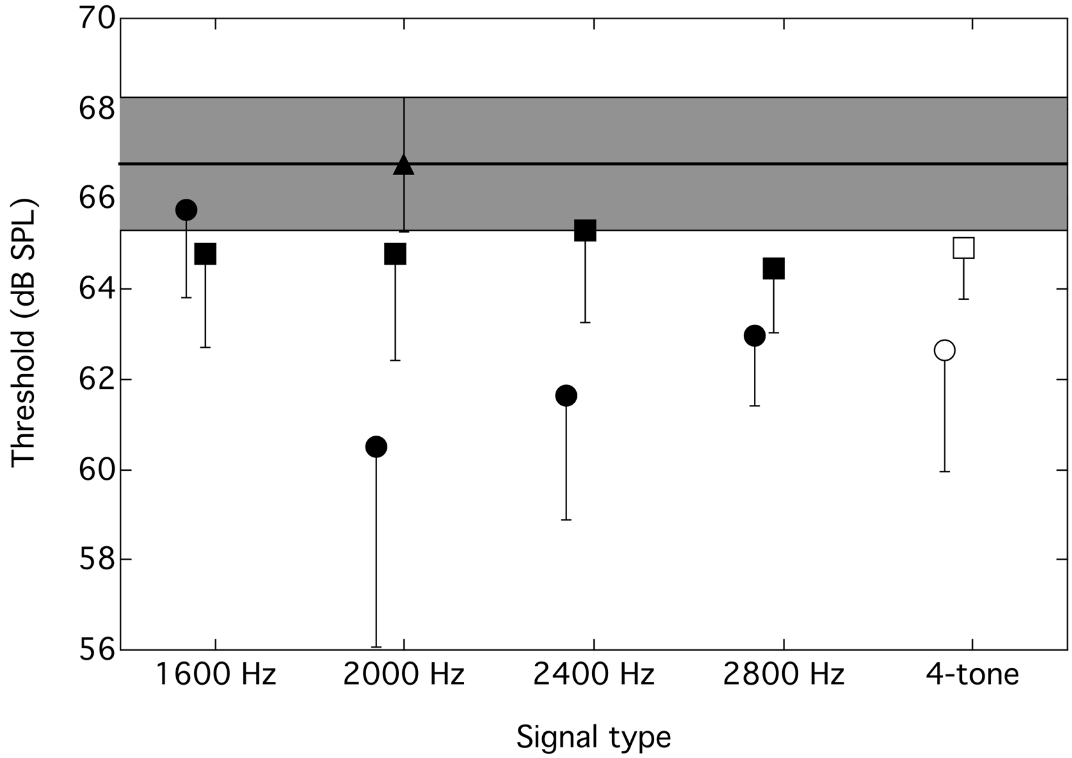

Figure 2.
Group mean detection thresholds for three masking conditions of experiment 1: 2000-Hz OSB (triangle & horizontal line, with hatched area = ±1 standard deviation); four-band comodulated masker (circles); 4-band comodulated masker with random temporal fringe (squares). Thresholds for the complex signal are highlighted as unfilled symbols. Error bars indicate −1 standard deviation.

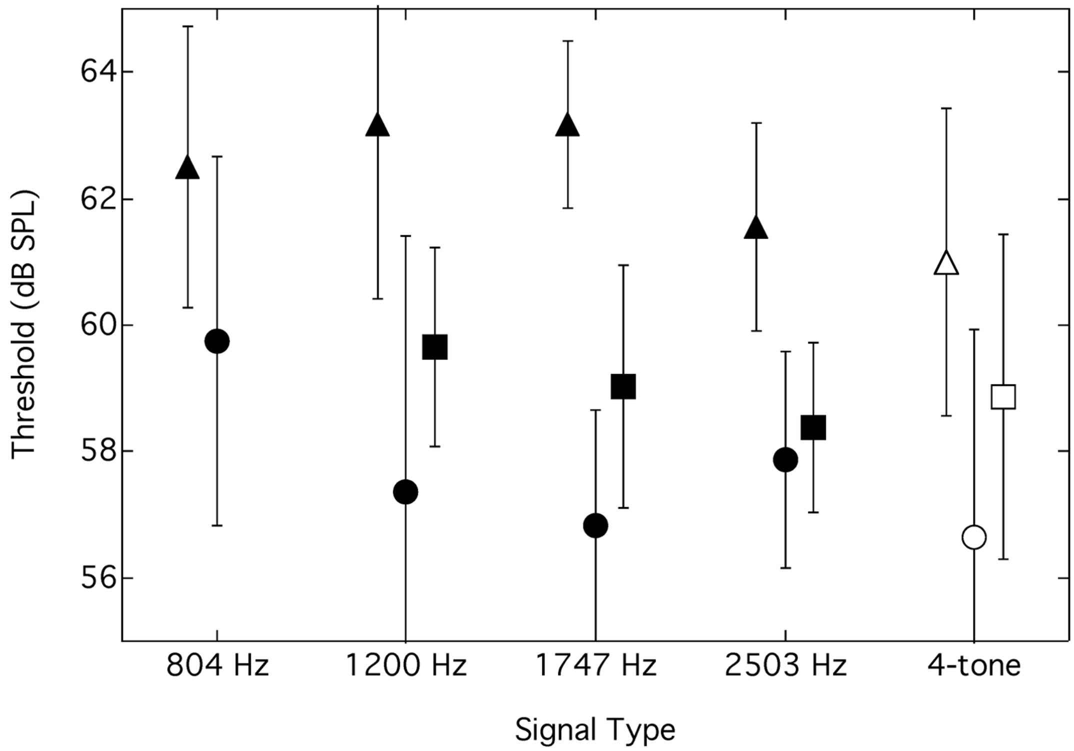

Figure 3.
Group mean detection thresholds for three masking conditions of experiment 2: four-band random masker (triangles); four-band comodulated masker (circles); 4-band comodulated masker with random temporal fringe (squares). Thresholds for the complex signal are highlighted as unfilled symbols. Error bars indicate ±1 standard deviation.

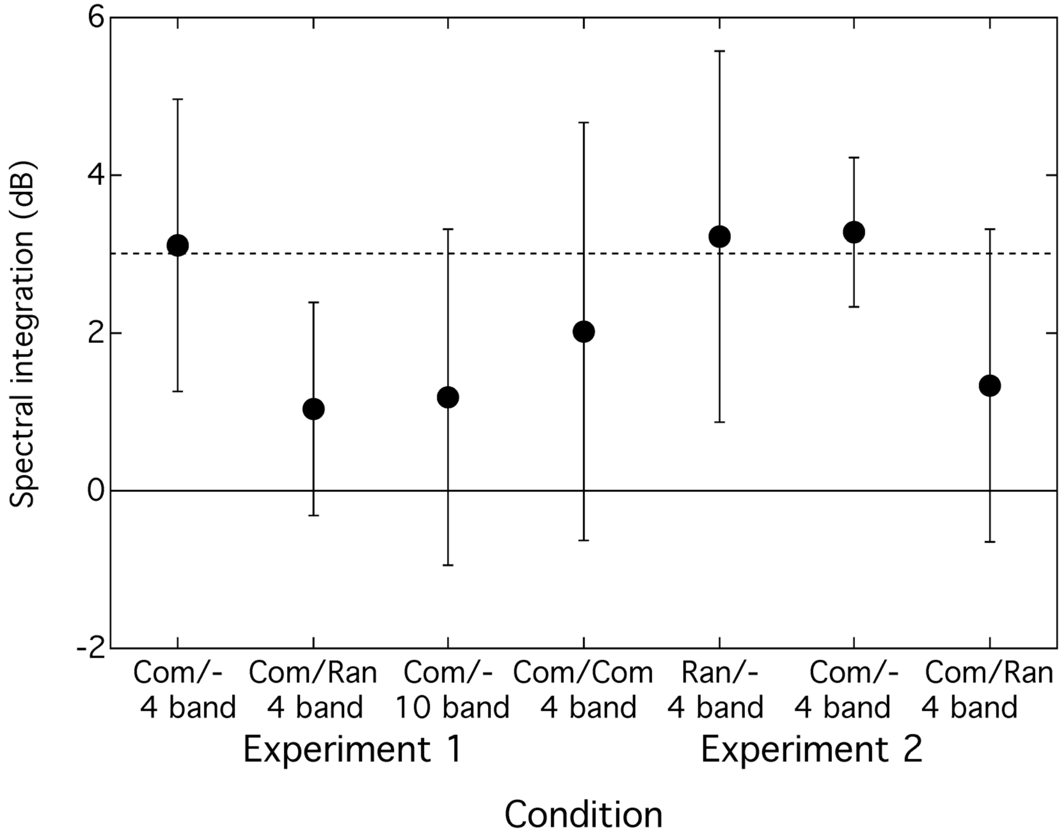

Figure 4.
Group mean magnitudes of spectral integration, ±1 standard deviation, for seven masking conditions. From Experiment 1, 4-band comodulated, gated (Com/−); 4-band comodulated with random fringe (Com/Ran); 10-band comodulated, gated (Com/−); and 4-band comodulated, continuous (Com/Com). From Experiment 2, 4-band random, gated (Ran/−); 4-band comodulated, gated (Com/−); 4-band comodulated with random surround (Com/Ran). Lower solid line references no spectral integration; upper dashed line references expected spectral integration using $10\log\sqrt{N}$ rule.