Indications for temporal fine structure contribution to co-modulation masking release

Robert H. Pierzycki[a)] and Bernhard U. Seeber
MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, United Kingdom

(Received 27 January 2009; revised 13 September 2010; accepted 14 September 2010)

The contribution of temporal fine structure (TFS) information to co-modulation masking release (CMR) was examined by comparing CMR obtained with unprocessed or vocoded stimuli. Tone thresholds were measured in the presence of a sinusoidally amplitude-modulated on-frequency band (OFB) of noise and zero, two, or four flanking bands (FBs) of noise whose envelopes were either co- or anti-modulated with the OFB envelope. Vocoding replaced the TFS of the tone and masker with unrelated TFS of noise or sinusoidal carriers. Maximum CMR of 11 dB was found as the difference between the co- and anti-modulated conditions for unprocessed stimuli. After vocoding, tone thresholds increased by 7 dB, and CMR was reduced to about 4 dB but remained significant. The magnitude of CMR was similar for both the sine and the noise vocoder. Co-modulation improved detection in the vocoded condition despite the absence of the tone-masker TFS interactions; thus CMR appears to be a robust mechanism based on across-frequency processing. TFS information appears to contribute to across-channel CMR since the magnitude of CMR was significantly reduced after vocoding. Since CMR was evidenced despite vocoding, it is hoped that co-modulation would also improve detection in cochlear-implant listening.
[DOI: 10.1121/1.3500673]

PACS number(s): 43.66.Dc, 43.66.Ts, 43.66.Mk, 43.66.Fe [RYL] Pages: 3614–3624

I. INTRODUCTION

Hearing-impaired listeners, including those with cochlear implants (CIs), demonstrate reduced ability to benefit from temporal fine structure (TFS) information (fast fluctuations over time with rates close to the signal's center frequency) (Nelson *et al*., 2003; Hopkins *et al*., 2008; Moore, 2008; Hopkins and Moore, 2009; Lorenzi *et al*., 2009). Lorenzi *et al*. (2006) measured consonant identification in quiet using stimuli processed such that the envelope information (slow amplitude fluctuations over time) was removed and the TFS information was preserved. They found that speech perception based on TFS cues is poor for hearing-impaired subjects in comparison with normal-hearing subjects. They also examined the role of envelope cues in speech perception using vocoder processing. A vocoder extracts the envelope in different frequency channels and uses it to modulate carrier signals like band-pass noises or tones (Dudley, 1939; Shannon *et al*., 1995). Identification of consonants in vocoded speech in quiet was almost as high as that of the unprocessed speech for both hearing-impaired and normal-hearing subjects (Lorenzi *et al*., 2006). Thus, ability to use envelope cues remains almost intact and leads to high intelligibility of speech in quiet.

However, envelope cues alone are not sufficient for hearing-impaired listeners to maintain good speech intelligibility when background sounds (maskers) are present. Both normal-hearing and hearing-impaired listeners suffer a reduction in speech intelligibility when a steady masker is present. When the masker is amplitude-modulated, a release from masking can be observed in the form of improved speech perception with normal-hearing listeners, but hearing-impaired listeners still perform poorly when maskers are modulated or fluctuating in level (Festen and Plomp, 1990; Nelson *et al*., 2003; Stickney *et al*., 2004; Lorenzi *et al*., 2006). The release from masking is normally attributed to listeners' ability to exploit dips in the modulation of the masking sound, i.e., it depends on envelope cues. However, recent findings show that masking release may also be affected by the ability to use TFS information (Füllgrabe *et al*., 2006; Lorenzi *et al*., 2006; Gnansia *et al*., 2008; Hopkins and Moore, 2009).

Speech understanding depends on multiple redundant cues. Some of the cues, such as voicing and manner, utilize both envelope and TFS information (Rosen, 1992). This makes it difficult to study the role of TFS in masking release and might be a reason for different outcomes between studies. For example, when the TFS information in the target speech was randomized by noise vocoder processing, Füllgrabe *et al*. (2006) and Nelson *et al*. (2003) found no release from masking. However, some masking release was observed after sine vocoder processing with TFS contributing about one third to its magnitude (Gnansia *et al*., 2008). The reason for no release from masking in the studies of Füllgrabe *et al*. (2006) and Nelson *et al*. (2003) could be the use of only four channels in their vocoders as opposed to 32 channels in Gnansia *et al*. (2008). Coarse frequency resolution might thus be detrimental to masking release with speech. Furthermore, masking release and contribution of TFS observed in the study of Gnansia *et al*. (2008) could indicate the importance of using sine carriers instead of noise carriers in vocoder processing. However, TFS apparently did not contribute to the large masking release found with a 32-channel sine vocoder in the study of Hopkins

[a)]Author to whom correspondence should be addressed. Electronic mail: rpierzycki@gmail.com

and Moore (2009). As shown by Zeng *et al.* (2005), the differences may also be partly due to the choice of speech material, which included vowel-consonant-vowel (VCV) utterances (Füllgrabe *et al.*, 2006; Gnansia *et al.*, 2008), IEEE sentences (Nelson *et al.*, 2003; Hopkins and Moore, 2009), or the Hearing in Noise Test (HINT) sentences (Qin and Oxenham, 2003). It can be concluded that at present it is difficult to obtain a consistent answer on the role of TFS information in speech understanding, suggesting that testing masking release for less complex stimuli than speech could possibly give a better indication of the role of TFS.

Detection thresholds of less complex stimuli like tones can be reduced when the masker carries coherent modulations across frequency channels (Hall *et al.*, 1984). This effect has been termed co-modulation masking release (CMR). In the flanking-band class of CMR experiments, a tone is masked by a single band of noise centered at the tone's frequency (on-frequency band, OFB) or by the OFB together with flanking bands (FBs) of noise. The FBs are presented at different frequency separations from the OFB, below or above the OFB. Thresholds of the OFB-masked tone are reduced when FBs with envelopes co-modulated (CM) with the OFB envelope are added to the OFB (Hall *et al.*, 1984, 1990; Schooneveldt and Moore, 1987, 1989).

CMR violates the critical band concept which states that masking energy outside a critical bandwidth corresponding to a single frequency channel (auditory filter) does not influence detection of a signal (Fletcher, 1940). Instead, it appears that the auditory system compares information across frequency channels to improve signal detection when the envelopes in those channels are co-modulated. The dip-listening model of CMR (Buus, 1985) postulates that the auditory system employs the masker's temporal minima across frequency channels as a cue for when to listen for the signal in the on-frequency channel. This model coincides with the dip-listening explanation for masking release in speech perception.

CMR can be partially explained by interactions of masker components within a single channel which makes it difficult to estimate across-channel CMR. Schooneveldt and Moore (1987) showed that CMR was largest when FBs were close in frequency to the OFB masker, and CMR decreased with increasing OFB and FB separation. Furthermore, CMR increased even further when stimuli were "transposed" to higher signal frequencies where auditory filters are wider in bandwidth. This suggests that CMR at close FB separations is affected by temporal interactions within the target channel, so called within-channel interactions, and should not be considered as a "true" across-channel CMR. A possible explanation for within-channel CMR is that at small spectral distances, interactions of the OFB and FB maskers can lead to temporal beats and provide a detection cue (McFadden, 1986; Schooneveldt and Moore, 1987). However, CMR can be observed for OFB and FB separations as large as three octaves (Cohen, 1991), which is likely a consequence of across-channel processing. Taken together, although CMR is generally seen as an envelope-based process, the contribution of TFS is not entirely clear on the basis of current evidence, particularly for the flanking-band paradigm where within-channel interactions are minimized through large separations of the FBs.

In the present experiments we used a flanking-band CMR paradigm with unprocessed or vocoded stimuli to examine the effect of the TFS on CMR. We argue that if across-channel processing relies solely on envelope information, a signal's TFS or signal-masker TFS interactions should not be required for CMR, unlike for masking release in speech (Füllgrabe *et al.*, 2006; Lorenzi *et al.*, 2006; Gnansia *et al.*, 2008). To examine effectiveness of listening in the dips, thresholds of a tone were measured in the presence of a 100% sinusoidally amplitude-modulated (SAM) noise masker. Thresholds were measured as a function of the number of FBs whose envelopes were either CM or anti-modulated (AntiM) with the OFB. We calculated CMR according to two definitions:

- CMR 1: The difference between the thresholds in the OFB masker alone condition and when CM FBs are present, CMR $1 = \text{OFB} - \text{CM}$, is positive when thresholds are reduced after adding FBs, an indicator for CMR,
- CMR 2: The threshold difference between the AntiM and CM conditions, CMR $2 = \text{AntiM} - \text{CM}$, is positive when thresholds are higher with AntiM compared to CM FBs, likewise indicating CMR.

The current vocoder approach offers an alternative to previous studies on the contribution of within-channel/TFS cues to CMR because it avoids the detection of increased pitch strength when the tonal signal is present in the background noise. The pitch of the tonal target provides a strong detection cue which could have been advantageous for CMR in previous experiments. However, when the stimuli are processed by the noise vocoder, the signal is not present as a spectral peak or in F0-periodic TFS. Instead, the signal is "represented" in the envelope of a one ERB_N-wide OFB signal [ERB_N: equivalent rectangular bandwidth of the auditory filter for normal-hearing listeners (Glasberg and Moore, 1990)]. More importantly, vocoding ensures that the TFS interactions of the signal and the OFB noise manifested in (1) the increased TFS periodicity of the noise due to the presence of the periodic TFS of the tone and (2) the phase change that occurs at the onset of the tone (Langhans and Kohlrausch, 1992) are not present in the TFS of vocoded stimuli and cannot be used as a detection cue. Therefore, in the vocoded experiments detection should only be performed on changes in the envelope.

Moreover, the vocoder paradigm can avoid issues connected with inferring the contribution of TFS to CMR from measurements with high-frequency stimuli. Based on the assumption that the lack of phase locking at higher frequencies prevents the use of TFS cues, a contribution of TFS cues to CMR was assumed from the decline of CMR at high frequencies (Moore, 2008). However, the limit of phase locking to TFS is not known in humans and recent studies show that phase-locked responses are still detectable at high frequencies (Recio-Spinoso *et al.*, 2005). Additionally, it is not clear if auditory processing would differ for CMR at different cochlear places. Auditory filters are wider at high frequencies and more FB energy could fall into the target channel, leading to an increase of within-channel interactions (Schooneveldt and Moore, 1987). Furthermore, using high-frequency OFB and FBs may affect the modulation spectrum because the stimuli are processed in wider auditory filters.

In addition to the OFB − CM definition of CMR, the present approach calculates CMR as the difference in thresholds obtained with CM and AntiM FBs. This has the advantage that the modulation spectrum is similar across the two comparison conditions with the main difference being the phase shift in the dominant SAM component. Furthermore, the depth of modulation is kept constant unlike when using random noise as a comparison to the CM condition. Also, modulators are predictable over time, potentially affecting grouping.

The experimental design presented here attempts to facilitate comparison with studies in electric hearing. A flanking-band paradigm with noise bands of one ERB_N bandwidth was used to mimic the roughly position-independent spread of excitation in electric hearing. Furthermore, using a flanking-band paradigm could help avoid issues with current spread that would occur in a band-widening CMR paradigm. Finally, the low pitch salience of the noise bands in our experiments with a noise-band vocoder gives a somewhat similar pitch perception to that with CIs (Laneau *et al.*, 2006).

Experiment 1 tested the effect of replacing the TFS of masker and probe by random TFS information on CMR with normal-hearing listeners by comparing CMR obtained with unprocessed against noise-vocoded stimuli. In experiment 2, a larger FB separation was used to assess the contribution of within-channel cues against across-channel processing in conditions with unprocessed stimuli. This was further tested by measuring masking from the FBs in the OFB channel. Finally, to establish whether the intrinsic modulation of the noise carriers affected the noise vocoder results, experiment 3 repeated conditions of experiment 2 with a sine vocoder.

II. EXPERIMENT 1: CMR WITH UNPROCESSED STIMULI AND A NOISE VOCODER

A. Methods

1. Conditions

The flanking-band CMR paradigm was used to measure thresholds of a tonal signal masked by the OFB alone (no FBs) or the OFB with added FBs. Either two or four FBs were presented in pairs around the OFB symmetrically on the ERB_N frequency scale (Glasberg and Moore, 1990). The top panel of Fig. 1 shows a sketch of the maskers and tone configurations in experiment 1. The FBs were either CM or AntiM, i.e., their envelope modulation was either in phase or π-shifted with respect to the OFB. Examples of the OFB and FB noise waveforms are shown in Fig. 2. To examine the effect of randomized TFS on CMR, signal thresholds were measured for unprocessed stimuli and those processed with a noise vocoder.

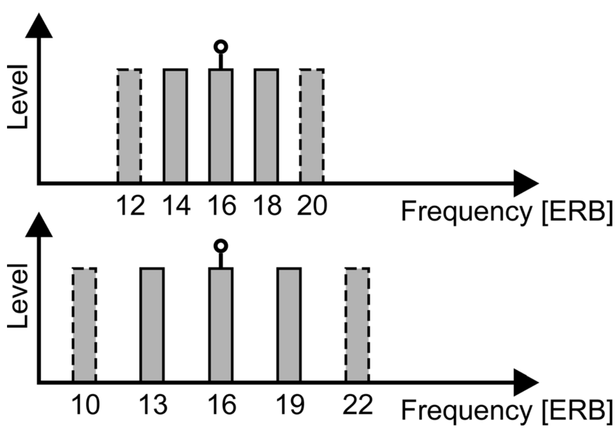

FIG. 1. Configuration of maskers and the tone in experiment 1 with 1 ERB_N (top panel) separation of FBs and of experiment 2 with 2 ERB_N (bottom panel) separation of FBs. The tone and the OFB masker were centered at the 16th ERB_N and FB pairs were added symmetrically on the ERB_N frequency scale to the OFB in different conditions.

2. Stimuli

Five frozen noise bands of 500-ms duration and a bandwidth of one ERB_N each were generated by restricting spectra of wideband, Gaussian noises to the desired ERB_N bandwidth

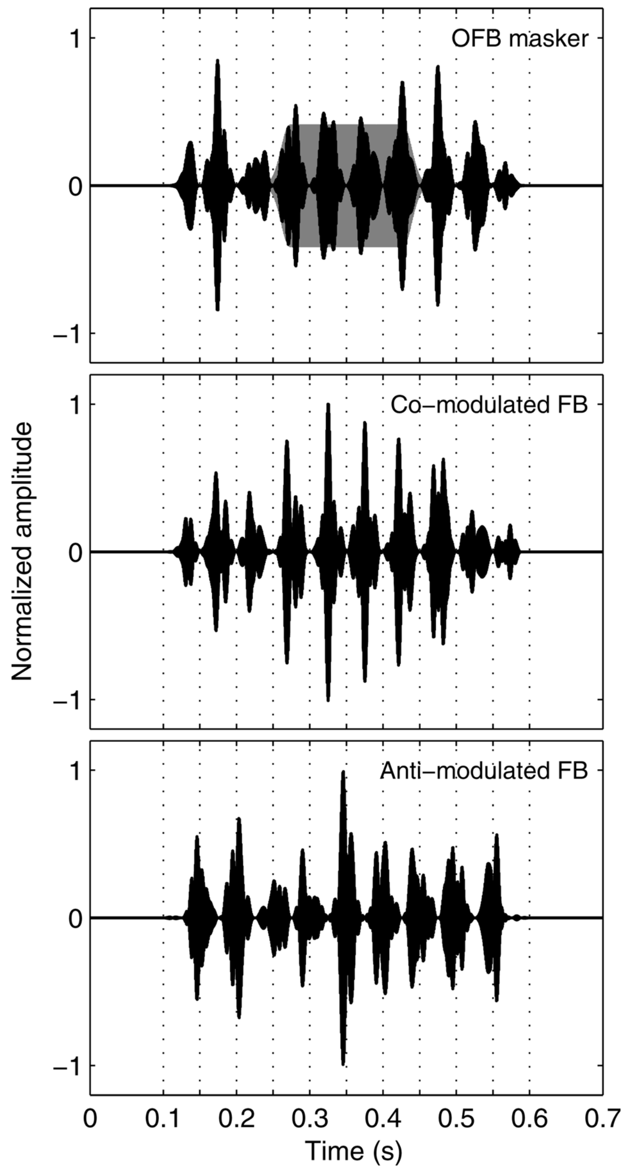

FIG. 2. Stimuli waveforms of the OFB (top panel), a CM (middle panel) and an AntiM FB (bottom panel). The envelope of the tone is indicated in the top panel (light grey). Vertical lines are plotted to visualise the co-modulation and the π-shift of the minima of the AntiM FB with respect to the OFB.

TABLE I. ERB_N number, center frequency (f_c), bandwidth, and lower (f_l) and upper (f_u) corner frequencies of masking noise bands used in experiments 1 and 2.

ERB_N	f_c (Hz)	Bandwidth (Hz)	f_l (Hz)	f_u (Hz)
10	443	72	407	479
12	605	89	560	649
13	699	100	649	749
14	805	111	749	860
16	1053	138	984	1122
18	1361	171	1275	1446
19	1541	190	1446	1636
20	1742	212	1636	1848
22	2216	263	2084	2347

and taking their inverse fast Fourier transform (FFT) to obtain noise bands in the time domain (FFT-filtering, 44.1-kHz sampling rate). The OFB noise band was centered at the frequency corresponding to the 16th ERB_N while the FBs were centered at ERB_N 12, 14, 18, and 20 (Glasberg and Moore, 1990), resulting in a gap between bands of 1 ERB_N (see Fig. 1, top panel). The center and corner frequencies of each noise band are listed in Table I. Each noise band was presented at 60 dB SPL.

Some studies have shown that using frozen as opposed to running noise maskers can affect detection thresholds (Langhans and Kohlrausch, 1992). However, our experiments show no significant difference between thresholds and CMR obtained with running vs frozen noise stimuli in the unprocessed case (see experiment 3). We decided to use frozen noise stimuli in the experiments with the noise vocoder since they allowed better control over the correlation of envelopes across presentation intervals.

The OFB and FBs were SAM at a modulation rate of 20 Hz and 100% modulation depth. Envelope maxima were unaffected by the modulation which results in a reduction of average, but not peak level. The modulation of the OFB started in its minimum, i.e., the phase was equal to $\varphi_m = -\pi/2$. The FBs were modulated either in phase with the OFB ($\varphi_m = -\pi/2$, CM) or in anti-phase with the OFB ($\varphi_m = +\pi/2$, AntiM). Modulated noise bands were gated with 20-ms Gaussian slopes.

A modulation frequency of 20 Hz was used to represent multiple modulation periods during the presentation of the stimuli to permit grouping of noise bands on the basis of common amplitude modulation (Grose and Hall, 1993). This frequency is within the range of modulation frequencies of speech (Rosen, 1992) and within the range of frequencies giving large release from masking in CMR experiments (Haggard *et al.*, 1990; Verhey *et al.*, 2003).

In the two FBs condition FBs were symmetrically placed around the OFB at 14 and 18 ERB_N while in the four FBs condition two outer FBs were added. The sinusoidal probe of 240-ms duration (20-ms Gaussian slopes) was placed at the center frequency of the OFB (16th $ERB_N \approx 1053$ Hz). The signal duration was chosen such that its onset and offset were in the dips of modulation of the noise maskers as shown in Fig. 2. All stimuli were generated using custom MATLAB software.

3. Vocoder processing

In the vocoded condition OFB and FB stimuli were filtered separately using a zero-phase (*filtfilt* function in MATLAB) tenth-order Butterworth band-pass filter. The filters were matched in center frequency and bandwidth to the corresponding noise bands at ERB_N 12, 14, 16, 18, and 20. The envelopes were extracted from the band-pass filtered signals by half-wave rectification and low-pass filtering at 200 Hz (zero-phase sixth-order Butterworth). Narrow-band noise carriers, matched in center frequency and bandwidth to the corresponding input filters, were modulated with the extracted envelopes. Each of the modulated narrow-band noises was then filtered using identical band-pass filters as in the input stage. The output signals were corrected to yield the same level per ERB_N as the original stimuli. Separately processed OFB and FBs were then added together depending on the test condition. To obtain the vocoded masker together with the probe, the OFB masker was summed with the tone before it was delivered to the band-pass filter at the frequency channel corresponding to the OFB. Input and output filters as well as carrier noise bands of the vocoder were thus matched in bandwidth and frequency to the OFB and FBs. This resulted in processing where the signals kept nearly the same bandwidth, level, and envelope, but received a new TFS unrelated to the original signal. Since the tone was added to the OFB before vocoding, the signal in the vocoded case was only represented in the envelope fluctuations of the OFB masker, while the TFS ("zero crossings") was unrelated to the signal. Detection could thus only be performed on changes in the envelope of the noise band across presentations, the reason why the same (frozen) noise carriers were used in target and non-target presentation intervals.

ERB_N-wide instead of constant bandwidth noise bands (Hall *et al.*, 1984; Schooneveldt and Moore, 1987) were chosen because: (1) The width over which neurons are stimulated from a single electrode in electric hearing is roughly independent of place. In normal hearing this is mirrored by auditory filters which also cover a roughly similar length on the cochlea while their absolute bandwidth increases with frequency. One motivation for the present work is to establish a paradigm that can be applied to electric hearing. To achieve comparability with electric hearing, noise bandwidth should be proportional to auditory filter width. Furthermore, due to the overlap of auditory filters, choosing constant bandwidth noises would stimulate a different number of filters in low and high-frequency regions, potentially leading to a different amount of CMR. Related to this Haggard *et al.* (1990) have shown that CMR magnitude is constant across center frequency when FBs are scaled in terms of auditory filter width. (2) To make the vocoder processing as "transparent" as possible, the carrier noise bandwidth follows that of the OFB and FBs.

4. Procedure

Subjects were seated in a soundproof booth and stimuli were presented monaurally to the left ear via Sennheiser HD 600 headphones. Stimuli were played from MATLAB via a digital soundcard (M-Audio Audiophile 2496). The digital sound signal was delivered to a custom made headphone amplifier

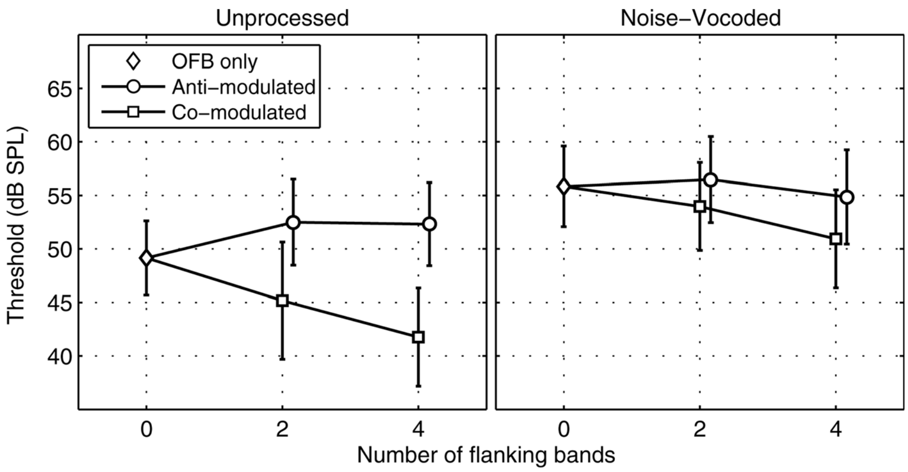

FIG. 3. Thresholds of the tone as a function of the number of FBs in experiment 1 with unprocessed (left) and noise-vocoded (right) stimuli. OFB and FB separation was 1 ERB_N. Results are presented as means across subjects' individual mean thresholds taken across three repetitions of each condition. The error bars represent plus and minus one standard deviation of the means across the subjects.

with built-in 24-bit D/A converter to which the headphone was connected.

A three-down/one-up adaptive, two-alternative forced-choice procedure was chosen to estimate the 79% point on the psychometric function (Levitt, 1971). Three stimulus intervals separated by 400-ms silence were played in each presentation. The first interval (anchor) always contained the masker, and the second and third intervals contained either the masker or the masker plus tone with equal probabilities. The vocoder used the same noise carriers in each of the three intervals of a presentation. Thus, changes in the extracted envelope due to the presence of the tone in one interval were reproduced in the vocoded signal and were not affected by the randomness of the carrier. Nevertheless, new random samples of noise were used across trials so that the potential interactions with the random inherent fluctuations in the noise carriers would not form a consistent cue. The subject's task was to specify whether the second or the third interval was different from the first one by pressing either "2" or "3" on a computer keyboard. The initial level of the signal was always above the expected threshold level and was equal to 70 dB SPL. The level was reduced after three consecutive correct responses and it was increased after each incorrect response, where the initial step of the level change was 6 dB, which was reduced to 4 dB after the first and to 2 dB after the second reversal. Twelve reversal points were collected and the threshold was calculated as the mean of the last eight reversals. A single run corresponded to a single, full response track. A single session consisted of a block of five runs. Each run in a session was derived for one of the following maskers: (1) OFB only, (2) OFB with two CM or (3) two AntiM FBs, and (4) OFB with four CM or (5) four AntiM FBs. The order of runs was random in each session. Unprocessed and vocoded conditions were presented in separate sessions. Subjects spent at least 20 min listening to the stimuli and familiarizing themselves with the experimental procedures, and completed three sessions for both the unprocessed and vocoded conditions over 3 days.

5. Subjects

Seven normal-hearing subjects, four male and three female, participated in the experiment. The subjects were between 21 and 32 yrs of age and had audiometric thresholds of 20 dB HL or less for frequencies from 250 to 8000 Hz in octave steps. Three of the subjects, including the first author, had prior experience in psychoacoustic experiments. Subjects received payment for their participation at a rate set by the Institute of Hearing Research. The experiment was approved by the Ethics Committee of the Psychology Department at the University of Nottingham.

B. Results

Thresholds of the tone in the unprocessed and vocoded conditions are shown in Fig. 3 as mean results with standard deviations across the means of the three trials collected for each subject and condition.[1] Table II shows derived CMR results.

We consider CMR 1, the OFB $-$ CM difference, first. For unprocessed stimuli, addition of CM FBs to the OFB masker leads to significant decrease of tone thresholds ([$F(2, 12) = 10.88$; $p < 0.003$], repeated measures analysis of variance (RM ANOVA); number of FBs as within-subject factor and three levels: 0, 2, and 4). Pair-wise comparisons after ANOVA revealed that that decrease is significant only when four FBs are present giving CMR 1 of about 7 dB ($p < 0.02$, Sidak correction). Vocoding of the stimuli leads to an overall increase of the CM thresholds ([$F(1, 6) = 256.99$; $p < 0.0001$], RM ANOVA; processing and FB number as factors with 2 and 3 levels, respectively). Nevertheless, CMR 1 of about 5 dB remains significant for four FBs ($p < 0.05$) despite the randomization of the TFS through vocoding.

TABLE II. CMR results as the mean across all subjects (experiments 1 and 2; frozen noise stimuli). CMR values were obtained from the difference of the means across three detection thresholds measured for the conditions indicated in column CMR.

		Unprocessed (dB)		Noise vocoder (dB)	
FBs separation	CMR[a]	Two FBs	Four FBs	Two FBs	Four FBs
1 ERB_N	1: OFB − CM	4	7*	2	5*
(experiment 1)	2: AntiM − CM	7**	11**	3**	4*
2 ERB_N	1: OFB − CM	4	6**	4	4
(experiment 2)	2: AntiM − CM	6**	8**	3**	4**

[a]RM ANOVA significance *0.05, **0.01.

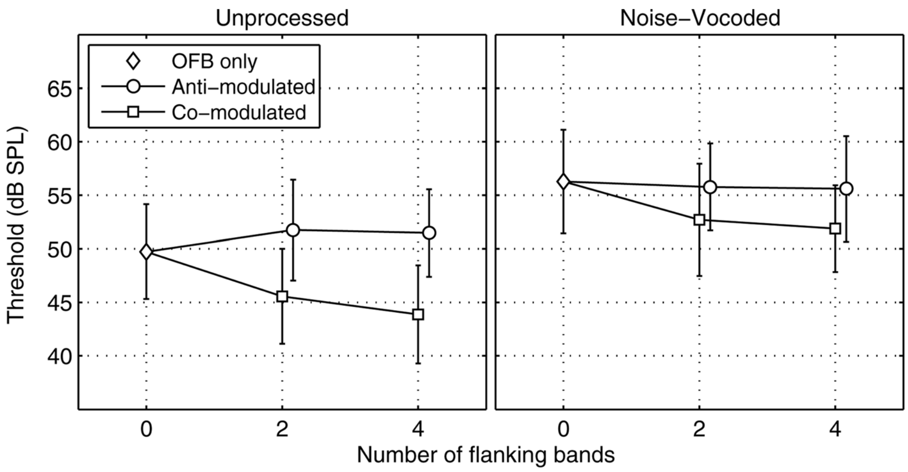

FIG. 4. Thresholds of the tone as a function of the number of FBs in experiment 2 with 2 ERB_N separation between FBs and the OFB. Results are presented as in Fig. 3 for unprocessed (left) and noise-vocoded (right) stimuli.

We now consider CMR 2, the difference AntiM − CM. Tone thresholds measured in the AntiM condition without processing are significantly higher than those measured for the CM condition when either two [$F(1, 6) = 30.75$; $p < 0.002$] or four FBs [$F(1, 6) = 38.00$; $p < 0.0009$] are present (RM ANOVA with factor modulation: CM and AntiM). Maximum CMR 2 is about 11 dB in this case (Table II). After vocoding thresholds of the AntiM FBs increase significantly ([$F(1, 6) = 22.34$; $p < 0.004$], RM ANOVA; processing and FB number as factors each with two levels). CMR 2 is found when either two [$F(1, 6) = 34.48$; $p < 0.002$] or four FBs [$F(1, 6) = 12.09$; $p < 0.02$] are added to the OFB masker and is largest, about 4 dB, with four FBs present.

Three two-way RM ANOVAs were used to estimate the effects of vocoding on the values of CMR 1 and CMR 2 (processing as the within-subject factor, two levels: Unprocessed and noise-vocoded). The decrease in CMR 1 observed after vocoding is not significant [$F(1, 6) = 3.71$; $p > 0.1$]. However, the reduction in CMR 2 of about 4 and 7 dB found after vocoding is significant for both two [$F(1, 6) = 16.19$; $p < 0.007$] and four FB [$F(1, 6) = 23.09$; $p < 0.003$] conditions, respectively.

Large CMR of up to 11 dB was observed in the unprocessed condition, but it was smaller after vocoding. In the first experiment FBs were separated from the OFB by a gap of 1 ERB_N to minimize the interaction between their components. However, the shape of auditory filters is not ideally rectangular as the filters have gradual slopes. It may occur that energy from the FBs spreads to the OFB filter and leads to within-channel interactions. These interactions are assumed to contribute to the magnitude of CMR at close frequency separations of the maskers (McFadden, 1986; Schooneveldt and Moore, 1987). The within-channel interactions cannot provide a detection cue in the vocoded condition due to the randomness of the noise carriers but they might have contributed to the apparent difference in CMR magnitude between unprocessed and vocoded conditions. To ensure that measured CMR is due to across-channel processing, experiment 2 used FBs separated by a gap of 2 ERB_N from each other and the OFB. Furthermore, experiment 2 can demonstrate that TFS cues do not contribute to CMR in the vocoded case if the CMR magnitude is unaffected by the separation of FBs.

III. EXPERIMENT 2: CMR FOR DISTANT FBs

A. Stimuli and procedures

The separation of FBs was set to two ERB_N in this experiment. Four new 500 ms-long, one ERB_N-wide noise bands centered at ERB_N 10, 13, 19, and 22 (Glasberg and Moore, 1990), were used as FBs (see bottom panel of Fig. 1 and Table I). The same noise band at ERB_N 16 was used for the OFB. The same vocoder processing strategy as in experiment 1 was used in the present experiment with the only difference that filter and carrier noise frequencies were matched to the new center and corner frequencies of the noise maskers. Procedures were identical to experiment 1 and the same seven subjects participated.

B. Results

Figure 4 shows tone thresholds in the unprocessed and vocoded conditions as across-subject mean results with standard deviations of the individual subject means. Table II lists the derived amounts of CMR.

In the unprocessed condition CMR 1 is significant with four FBs present ($p < 0.004$) and amounts to about 6 dB. The thresholds increase significantly after vocoding [$F(1, 6) = 318.99$; $p < 0.0001$] and CMR 1 is not significant with either two ($p > 0.2$) or four ($p > 0.1$) FBs present.

In the unprocessed case, the AntiM and CM conditions were significantly different with two ($p < 0.0007$) or four FBs ($p < 0.0009$) present, giving maximum CMR 2 of about 8 dB with four FBs added. After vocoding the thresholds generally increase [$F(1, 6) = 39.68$; $p < 0.0008$], but the difference between the AntiM and CM conditions, expressing CMR 2, remains significant with both two [$F(1, 6) = 19.93$; $p < 0.005$] and four FBs [$F(1, 6) = 18.04$; $p < 0.006$] present. The addition of four FBs leads to maximum CMR 2 of about 4 dB.

Two-way RM ANOVAs were again used to assess the effect of vocoding on magnitude of CMR 2. The ANOVA results were similar to those obtained in experiment 1. Vocoding significantly reduces the magnitude of CMR 2 by about 3 and 4 dB when two [$F(1, 6) = 8.89$; $p < 0.03$] and four FBs [$F(1, 6) = 7.54$; $p < 0.04$] are present, respectively.

ANOVAs were also used to test whether increasing the FB separation from 1 to 2 ERB_N had a significant effect on CMR values (experiment 1 vs 2). Separate ANOVAs were used for the unprocessed and vocoded case. Two-way ANOVAs were run for the CMR 1 and CMR 2 conditions, respectively, with FB separation as factors (levels: 1 and 2 ERB_N). The analysis showed that the increase of FB separation to two ERB_N does not have a significant effect on the magnitude of CMR 1 [$F(1, 6) = 1.45$; $p > 0.2$], or the magnitude of CMR 2 with both two [$F(1, 6) = 0.69$; $p > 0.4$] and four FBs [$F(1, 6) = 4.72$; $p > 0.07$] in the unprocessed condition. The increase of FB separation is also not significant for CMR 2 in the vocoded condition with either two [$F(1, 6) = 0.63$; $p > 0.4$] or four FBs present [$F(1, 6) = 0.06$; $p > 0.8$].

C. Masking from FBs

The excitation pattern at the output of the auditory filters for a narrow band of noise such as used in the present study is not symmetric, but has a shallow slope on the high-frequency side (Fastl and Zwicker, 2007). This is due to the shallow low-frequency slope of auditory filters. A filter may thus reach far beyond its nominal bandwidth of 1 ERB_N, with energy particularly from lower frequencies contributing to the filter output. Thus, the FBs may contribute to masking in the OFB channel.

To examine masking from the FBs in the OFB channel, we measured thresholds of the vocoded tone masked by two vocoded FBs placed symmetrically around the tone (OFB), but without the OFB masker present. The ERB_N-wide FBs were separated from the OFB channel by a gap of 1 or 2 ERB_N, so that the spectral configuration of the FBs and the signal was the same as in experiment 1 or 2 (cf. Fig. 1 top and bottom, and Table I). However, FBs were not modulated, and the level of the FBs was equal to the peak level of the modulated FBs in experiments 1 and 2. This was done to achieve maximum masking, i.e., to estimate the worst case of energetic masking which occurs when the intensity of the flankers reaches its temporal maximum.

The FBs and the tone were vocoded using the same processing as in experiment 1 and 2. Only the tone was delivered to the OFB channel input of the vocoder, so that the tone's TFS was replaced by an ERB_N-wide noise equal in center frequency and bandwidth to the OFB, and multiplied by the extracted tone's envelope. This mimics the case in experiments 1 and 2 when the tone dominates the envelope of the vocoder output during the modulation minimum of the OFB.

The experimental procedure was the same as in experiment 1 and 2 and three of the seven subjects participated. In summary, the present masking experiment was identical to the vocoded conditions with two FBs of experiments 1 and 2, except that the OFB masker was not present and FBs were not modulated.

The threshold of the vocoded tone masked by two FBs alone is equal to about 36 dB SPL when the FBs are 1 ERB_N away from the OFB channel. When the FBs are 2 ERB_N away from the OFB channel the threshold decreases to about 25 dB SPL. These thresholds are 20 and 31 dB lower than thresholds in the vocoded OFB condition of experiment 1 and 2, respectively. We conclude therefore that the tone is dominantly masked by the OFB masker.

IV. EXPERIMENT 3: CMR WITH A SINE VOCODER

In experiment 2, the contribution of TFS to CMR was tested with a noise vocoder. However, the narrow-band noise carriers of the vocoder exhibit an inherent modulation which could affect CMR in two ways: (1) The extracted envelope is not exactly reproduced in the vocoded signal and (2) the envelope correlation between the OFB and FBs could be reduced compared to the unprocessed condition, potentially reducing CMR. To assure that the reduction in CMR magnitude after vocoding is not due to the fluctuations of the noise carriers in the vocoder, conditions of experiment 2 were repeated using a vocoder with tonal carriers whose envelopes are constant. Since the tonal carriers' envelopes are constant, experiment 3 was carried out with running noise stimuli.

A. Stimuli and procedures

Experiment 3 repeated the condition with 2 ERB_N FB separation and four FBs of experiment 2 with a sinusoidal vocoder. This condition should give maximum CMR while minimizing channel interactions.

Band-pass filtering and envelope extraction were identical to the vocoder used in experiment 2. However, the extracted envelopes were used to modulate tonal carriers with random starting phase at the center frequencies of the noise bands/analysis filters (Table I). As before, output filters identical to the band-pass filters of the input stage were used to limit spectral spreading to the width of a critical band. Channel levels were corrected to yield the same level per ERB_N as the input FB/OFB. In summary, vocoder processing was identical to experiment 2 except that tonal carriers placed at the band center frequencies were used instead of noise bands.

Stimuli were the same as in experiment 2 apart from the fact that new input noise bands were generated in each trial. This was done because tonal carriers of the vocoder have no inherent modulations and thus testing would have otherwise used the same stimuli repeatedly. Nevertheless, input noises were frozen across target and non-target intervals like in experiments 1 and 2 to assure the envelope changes stem only from the addition of the tone. Because of these alterations data were also obtained for the unprocessed conditions.

Procedures were identical to those in experiments 1 and 2. Seven subjects participated in experiment 3, three of them also contributed to the previous experiments, and all had normal audiometric thresholds.

B. Results

We first consider the effect of using running noise (experiment 3) instead of frozen noise (experiments 1 and 2) on thresholds in CM and AntiM conditions in the absence of vocoder processing. We used a mixed model RM ANOVA on CM

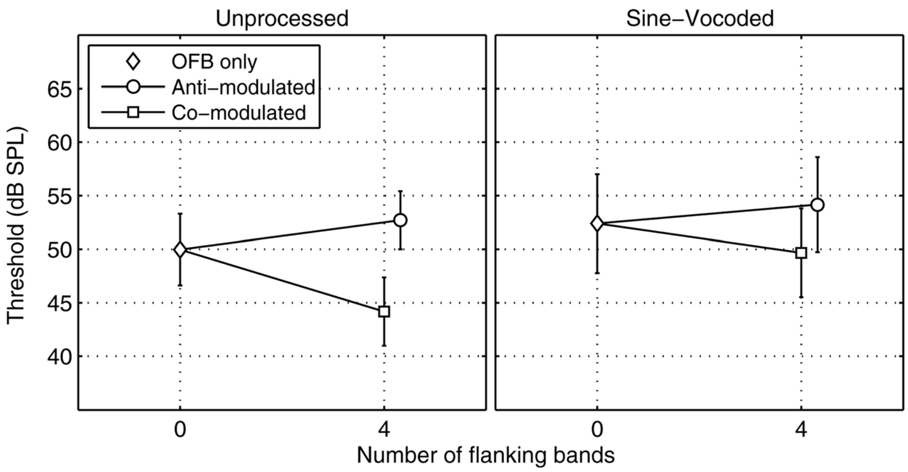

FIG. 5. Tone thresholds for unprocessed (left) and sine-vocoded stimuli for zero or four FBs present (experiment 3). Running noise was used in both conditions. Separation between FBs and the OFB was 2 $\mathrm{ERB_N}$. Results are presented as in Figs. 3 and 4.

and AntiM thresholds. ANOVA results showed no significant effect of noise type on thresholds in CM conditions ([$F(1, 5) = 0.15$, $p > 0.7$]; within-subject factors: Noise type with levels frozen and running, FB with levels 0 and 4; between-subject factor: Subject with levels same and different). There was also no significant effect of noise type on AntiM thresholds ([$F(1, 5) = 0.77$, $p > 0.4$]; noise type as within-subject factor and subject as between-subject factor; same levels as for the CM thresholds). Furthermore, there was no significant effect of subject for both CM [$F(1, 5) = 0.04$; $p > 0.8$] and AntiM [$F(1, 5) = 0.31$; $p > 0.6$] thresholds. Mixed RM ANOVAs were also used to test the effect of noise type on values of CMR 1 and CMR 2 between unprocessed conditions with frozen and running noise (noise type as within-subject factor plus subject as between-subject factor, same levels as in the ANOVA on thresholds). There was no significant effect of noise type on CMR 1 [$F(1, 5) = 0.06$; $p > 0.8$] and on CMR 2 [$F(1, 5) = 0.36$; $p > 0.5$], and no significant effect of subject on CMR 1 [$F(1, 5) = 0.03$; $p > 0.8$] and CMR 2 [$F(1, 5) = 0.09$; $p > 0.7$].

Figure 5 shows tone thresholds for unprocessed and sine-vocoded stimuli. In the unprocessed condition the addition of four CM FBs reduces thresholds by about 6 dB compared to the OFB only condition and by about 9 dB as compared to the AntiM condition giving significant CMR 1 [$F(1, 6) = 36.77$; $p < 0.001$] and CMR 2 [$F(1, 6) = 251.18$; $p < 0.0001$] (Table III). These CMR values are similar to those obtained in the corresponding conditions with frozen noise stimuli (cf. Table II). The threshold difference between the CM and OFB only conditions, and thus CMR 1, is not significant [$F(1, 6) = 3.11$; $p > 0.1$] when the stimuli are processed with a sine vocoder, which is similar to the corresponding conditions with noise-vocoded stimuli. However, thresholds differ significantly between AntiM and CM conditions [$F(1, 6) = 26.29$; $p < 0.003$], giving maximum CMR 2 of about 5 dB.

Results for the unprocessed case with 2 $\mathrm{ERB_N}$ FB separation and the sine-vocoded case were subjected to ANOVAs to test for the effect of processing as before. Sine vocoding significantly increases CM thresholds [$F(1, 6) = 41.57$; $p < 0.0007$], but does not have a significant effect on AntiM condition thresholds [$F(1, 6) = 2.45$; $p > 0.1$]. However, the difference of 4 dB in CMR 2 found between unprocessed and sine-vocoded conditions is significant [$F(1, 6) = 55.22$; $p < 0.0004$] and similar to that with the noise vocoder (cf. Tables II and III).

TABLE III. CMR results as the mean across all subjects (experiment 3; running noise stimuli). CMR values were obtained the same way as for Table II.

FBs separation	CMR[a]	Unprocessed (dB) Four FBs	Sine vocoder (dB) Four FBs
2 $\mathrm{ERB_N}$ (experiment 3)	1: OFB − CM	6**	3
	2: AntiM − CM	9**	5**

[a]RM ANOVA significance *0.05, **0.01.

V. DISCUSSION

The present study investigated the contribution of the signal-masker TFS interactions to CMR by comparing the magnitude of CMR obtained with unprocessed and vocoded stimuli. CMR 2, obtained as the difference between the CM and AntiM FBs, is observed consistently across all experimental conditions for stimuli that were either unprocessed or processed using the noise or sine vocoder. After randomization of the TFS cues through vocoder processing CMR 2 is reduced by 50%, independent of processing type (noise vs sine vocoder) or separation of FBs.

A. Contribution of the TFS cues to CMR

Fletcher (1940) assumed that masker energy outside the critical band does not affect signal detection, such that a single auditory filter centered at a signal's center frequency is used for detection. In the present experiments we introduced a number of factors to separate energetic masking from the CMR effect. The bandwidth of the maskers was fixed to one $\mathrm{ERB_N}$ to reduce any effect of masker bandwidth on detection within that critical band. To ensure equal potential contribution of all FBs, they were given an identical level per $\mathrm{ERB_N}$ (cf. uniform masking noise, Fastl and Zwicker, 2007). Finally, and most importantly, to limit spreading of FB energy into the OFB channel, noises had steep slopes and the

FBs were separated from the OFB by a gap of at least one ERB_N.

Although the experimental design limited energy spread across auditory filters, within-channel interactions could still affect CMR since "real" auditory filters have non-ideal slopes. Interactions can occur when energy spreads from neighboring FBs into the OFB channel, leading to temporal beats due to TFS interactions (McFadden, 1986; Schooneveldt and Moore, 1987). This may have affected detection in the unprocessed condition since beating could occur against the steady tone. However, within-channel interactions are unlikely to affect thresholds in the noise-vocoded condition for two reasons. First, the TFS of both the OFB signal and the FBs is random. Summing two random signals in the same filter results in a random signal with changed TFS, envelope, and intensity. For a within-channel interaction to contribute to CMR, the random signal must have properties that consistently lead to lower tone thresholds when summing occurs which is unlikely because of the randomness of the signals and because we used new carrier noises in every trial. Second, the tone is not represented in the TFS of the vocoded signal. Thus, it cannot alter the TFS of the masker nor can within-channel interactions from the FBs affect the TFS of the tone.

Experiment 2 showed that masking from the FBs is at least 20 dB below that from the OFB masker, demonstrating that the OFB channel is strongly dominated by the OFB masker. It is thus unlikely that the presence of the FBs could alter the envelope/TFS in the OFB to give a consistent detection advantage over the random fluctuation of the noise carrier. Additionally, changing the separation of the FBs from 1 to 2 ERB_N did not result in a significant change of CMR. Within-channel interactions are also unlikely to affect thresholds in the sine-vocoded condition because of the spectral separation to the FBs and because the OFB TFS does not reflect the TFS of the unprocessed signal. For these reasons we conclude that it is unlikely that within-channel effects contributed to CMR, particularly in the vocoded conditions. Assuming that within-channel interactions were negligible in the vocoded condition, CMR obtained with vocoding reflects the contribution of across-channel processing.

A comparison of the unprocessed and vocoded conditions revealed that not representing signal-masker interactions in the TFS leads to an overall increase of the detection threshold of the tone. This could be partially explained by the fact that noise must be detected in noise in the noise-vocoded condition while the unprocessed case is a tone-in-noise detection study. This could lead to a threshold difference of about 2–3 dB according to measurements of tone detection in noise and intensity discrimination for that noise (Bos and de Boer, 1966; Grose and Hall, 1997). The reason for the discrepancy between this 3-dB difference and the 7-dB difference observed here is not entirely clear, but it could be accounted for by the difference in stimuli and/or procedures used in previous studies. Bos and de Boer (1966) measured pure tone detection or noise intensity discrimination against continuous narrow-band noise maskers while Grose and Hall (1997) used maskers comprising of one to eight spectrally separated 20-Hz-wide noise bands presented together with the signal and a continuous broadband background noise. However, variability as large as 8 dB was found by Fantini *et al.* (1993) between tone-in-noise and noise-in-noise detection in the OFB masking condition of their CMR experiment which used gated maskers. A factor contributing to the above differences could be the lack of a strong pitch cue after randomization of the tone's TFS since, unlike in the magnitude spectrum of the unprocessed stimulus, there is no peak in the magnitude spectrum of the vocoded stimulus when the signal is added. This could lead to a reduction of segregation cues between signal and masker in the noise-vocoded condition (Cooke and Ellis, 2001).

Addition of CM FBs leads to CMR in the vocoded conditions even though overall thresholds are increased. After vocoding CMR cannot employ consistent TFS information as it has been replaced and is meaningless. Although this shows the strong contribution of envelope information to CMR, it nevertheless appears that TFS may also contribute to the magnitude of across-channel CMR since both noise and sine vocoding decreases it. This is surprising because previous studies have demonstrated a lack of sensitivity to phase (TFS) across frequency channels in the auditory periphery (Zwicker, 1952). However, the reduction of CMR with random TFS is in line with results on the perception of vocoded speech in amplitude-modulated maskers (see below).

B. Effect of vocoder processing

It could be argued that the envelope correlation across noise bands is affected by using uncorrelated carrier noises of different bandwidths, thereby making the comparison of unprocessed and noise-vocoded conditions more difficult. First, using uncorrelated noise carriers in noise vocoding reduces the across-channel correlation of envelopes from about 0.7 in the unprocessed conditions to about 0.6, but correlation remains almost unchanged for sine-vocoded stimuli because the sine vocoder with tonal carriers almost perfectly reproduces the noise bands' envelopes.[2] This could lead to reduced CMR magnitudes after processing with a noise vocoder as normal-hearing listeners were shown to be able to detect relatively small changes in envelope correlation (Richards, 1987). However, our results appear not to show such an effect of reduced correlation on CMR since the magnitude of CMR drops by the same amount and is almost identical for sine and noise vocoding despite observed changes in envelope correlation. Second, results could be affected by the modulation frequencies "added" by the noise vocoder due to intrinsic envelope modulations of the carriers, but these modulations are not correlated across noise bands. However, the correlation of slow rate envelope fluctuations of the applied SAM is not affected by vocoding, and this is what mainly contributes to the magnitude of CMR (Haggard *et al.*, 1990; Eddins and Wright, 1994). This can be accounted for by the 20-Hz modulation component having on average a 15-dB higher level than other components in the modulation spectrum, as found across bands as well as before and after vocoding. Furthermore, low modulation frequencies tend to contribute more strongly to the magnitude of CMR than higher modulation frequencies (Eddins and

Wright, 1994) while only the higher modulation frequencies are affected by the widening of the filters at higher frequencies. It appears therefore that any decrease of CMR in the vocoded conditions cannot be solely explained by changes in envelope correlation or modulation spectrum due to vocoding.

Vocoder processing of modulated maskers can also reduce the depth of modulation which in turn can reduce CMR without precautions taken (Verhey *et al.*, 1999). In the present study, band-limited maskers were modulated at a low rate (20 Hz) while the envelope extraction utilized a low-pass filter with a cut-off frequency ten times the modulation rate (200 Hz) to minimize any effects of vocoding on modulation depth. Analysis of the noise-vocoded stimuli showed that the modulation minima were on average about 20 dB below the threshold levels in the unprocessed OFB only condition. Because tone thresholds were far above the modulation minima it is unlikely that the reduction in CMR was due to a reduction in modulation depth. Thus, temporal masking from the modulation maxima was likely the factor controlling the "effective" modulation depth in the present experiments (Fastl and Zwicker, 2007) and a slight decrease of modulation depth due to vocoding should not affect the magnitude of CMR. Instead, the decrease in CMR after vocoding appears to be mainly connected with the lack of information in the TFS, in agreement with recent studies of masking release in speech.

C. Comparison with vocoded speech experiments

Significant, although reduced, CMR was observed for vocoded stimuli in the present study while no or little release from masking has been observed for vocoded speech (Qin and Oxenham, 2003; Füllgrabe *et al.*, 2006; Lorenzi *et al.*, 2006; Gnansia *et al.*, 2008). One of the reasons for such discrepancy could be that experiments with vocoded speech test speech perception based on identification which requires more information from stimuli than would be needed for their detection. Furthermore, TFS does not contribute independently of the temporal envelope to the information necessary for speech understanding, that is to voicing and manner (Rosen, 1992). Depending on the test material and the background sounds those features would have varying contribution to speech perception which renders interpretation of overall performance more difficult. Thus, we argue that detection of a tone in a flanking-band CMR experiment with vocoding could provide a more controlled way of testing the TFS contribution to the perception of masked signals.

Another reason for any discrepancies between the present results and those from masking release in speech experiments could be due to the vocoding strategies. Experiments with vocoded speech often use broadband maskers which can give rise to within-channel effects due to partially overlapping spectral channels or excitation (Qin and Oxenham, 2003; Füllgrabe *et al.*, 2006; Lorenzi *et al.*, 2006; Gnansia *et al.*, 2008). Here, clearly separated bands of noise were used to minimize within-channel effects and spread of energy across channels.

D. Role of grouping in CMR

It has been shown that the CMR mechanism does not act independently from auditory object formation (Grose and Hall, 1993; Dau *et al.*, 2005). In our experiment larger CMR was found for a larger number of FBs which is in line with previous findings (Hall *et al.*, 1990). This could be due to stronger grouping of the FBs and the OFB on the basis of common amplitude modulation. A larger number of FBs could give the listener clearer information for when to listen for the signal in the modulated masker (Buus, 1985). We argue that grouping by common amplitude modulation is still present after vocoding and propose it as a possible mechanism for perceptual "selection" of envelope information across auditory filters to aid detection in the OFB channel. This grouping process should function identically in the unprocessed and vocoded condition since the AntiM − CM difference was significant in all vocoded conditions proving that CMR is mostly based on envelope correlation at the outputs of auditory filters.

E. Implications for CIs

Most current CI processors replace the TFS of input signals by fixed rate pulse trains modulated with envelopes extracted in separate frequency channels. The CI literature has so far shown only the negative effect of applying multiple modulations in the form of modulation detection interference (Chatterjee, 2003). The present study instead demonstrates a positive effect of modulation in the form of CMR. Thus, present results suggest that across-frequency processes responsible for CMR could be active in listening with CIs even with some of the current processing strategies and contribute toward a benefit in certain situations. The results also show that the provision of TFS information may lead to an additional benefit as suggested in other studies (Lorenzi *et al.*, 2006).

VI. CONCLUSIONS

Contribution of TFS information to CMR was studied using a flanking-band CMR paradigm with unprocessed and vocoded stimuli. It was shown that co-modulation of the FBs and the OFB gives significant CMR in the unprocessed case. CMR was larger when more CM FBs were present, even for increased frequency separations of the FBs. Vocoding of the stimuli led to an overall increase of tone thresholds, which could be due to detection of noise-in-noise in the noise-vocoded condition as opposed to the detection of tone-in-noise in the unprocessed condition. This causes absence of a strong pitch cue, leading to poorer signal and masker segregation in the noise-vocoded condition. However, CMR was still observed after the tone's TFS was replaced by random TFS in the noise vocoder, though its magnitude was significantly reduced in comparison with CMR in the unprocessed case. Similar results were found for a sine vocoder which replaced the TFS of the FB and OFB signals by modulated tones. The reduction in CMR magnitude appears to be connected with the lack of meaningful TFS information even though the paradigm was designed to reduce the effect of within-channel cues and separate masking experiments showed that the contribution of the FBs to masking the tone in the OFB channel was far below that of the OFB masker. We conclude that envelope cues were predominantly responsible for CMR in the vocoded condition. We suggest CMR is partly based on dip-listening aided by grouping of the FBs by common envelope while the presence of consistent TFS leads to a further increase of CMR. Since CMR is observed even in the absence of the signal's TFS, we argue that this process should be available to CI listeners to improve signal detection in co-modulated backgrounds.

ACKNOWLEDGMENTS

This work was supported by the intramural program of the Medical Research Council (MRC, UK). The authors would like to thank Dr. Trevor Shackleton for helpful comments and suggestions on an earlier version of the manuscript. The authors are also grateful to Dr. Ruth Litovsky and three anonymous reviewers for their valuable comments and suggestions.

[1]Before calculating the mean for each subject across three runs all collected data were inspected for outliers. Tracks/trials for which the level range across the eight final reversals exceeded 15 dB were rejected.

[2]To measure correlation envelopes were extracted in the same way as in the vocoder processing, either directly from the unprocessed and vocoded stimuli or after filtering these stimuli with gammatone filters. There was no noticeable effect of filter shape (rectangular, Gammatone) on correlation.

Bos, C. E., and de Boer, E. (**1966**). "Masking and discrimination," J. Acoust. Soc. Am. **39**, 708–715.

Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Chatterjee, M. (**2003**). "Modulation masking in cochlear implant listeners: Envelope versus tonotopic components," J. Acoust. Soc. Am. **113**, 2042–2053.

Cohen, M. F. (**1991**). "Comodulation masking release over a three octave range," J. Acoust. Soc. Am. **90**, 1381–1384.

Cooke, M., and Ellis, D. P. W. (**2001**). "The auditory organization of speech and other sources in listeners and computational models," Speech Commun. **35**, 141–177.

Dau, T., Ewert, S. D., and Oxenham, A. J. (**2005**). "Effects of concurrent and sequential streaming in comodulation masking release," in *Auditory Signal Processing: Physiology, Psychoacoustics, and Models*, edited by D. Pressnitzer, A. de Cheveigné, S. McAdams, and L. Collet (Springer-Verlag, New York), pp. 335–343.

Dudley, H. (**1939**). "Remaking speech," J. Acoust. Soc. Am. **11**, 169–177.

Eddins, D. A., and Wright, B. A., (**1994**). "Comodulation masking release for single and multiple rates of envelope fluctuation," J. Acoust. Soc. Am. **96**, 3432–3442.

Fantini, D. A., Moore, B. C. J., and Schooneveldt, G. P. (**1993**). "Comodulation masking release as a function of type of signal, gated or continuous masking, monaural or dichotic presentation of flanking bands, and center frequency," J. Acoust. Soc. Am. **93**, 2106–2115.

Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics. Facts and Models* (Springer-Verlag, Berlin, Heidelberg), Chaps. 4, 6, 10, pp. 61–110, 149–174, 247–256.

Festen, J. M., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**, 1725–1736.

Fletcher, H. (**1940**). "Auditory Patterns," Rev. Mod. Phys. **12**, 47.

Füllgrabe, C., Berthommier, F., and Lorenzi, C. (**2006**). "Masking release for consonant features in temporally fluctuating background noise," Hear. Res. **211**, 74–84.

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

Gnansia, D., Jourdes, V., and Lorenzi, C. (**2008**). "Effect of masker modulation depth on speech masking release," Hear. Res. **239**, 60–68.

Grose, J. H., and Hall, J. W. III (**1993**). "Comodulation masking release: Is comodulation sufficient?," J. Acoust. Soc. Am. **93**, 2896–2902.

Grose, J. H., and Hall, J. W. III (**1997**). "Multiband detection of energy fluctuations," J. Acoust. Soc. Am. **102**, 1088–1096.

Haggard, M. P., Hall, J. W. III, and Grose, J. H. (**1990**). "Comodulation masking release as a function of bandwidth and test frequency," J. Acoust. Soc. Am. **88**, 113–118.

Hall, J. W. III, Grose, J. H., and Haggard, M. P. (**1990**). "Effects of flanking band proximity, number, and modulation pattern on comodulation masking release," J. Acoust. Soc. Am. **87**, 269–283.

Hall, J. W., Haggard, M. P., and Fernandes, M. A. (**1984**). "Detection in noise by spectro-temporal pattern analysis," J. Acoust. Soc. Am. **76**, 50–56.

Hopkins, K., and Moore, B. C. J. (**2009**). "The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise," J. Acoust. Soc. Am. **125**, 442–446.

Hopkins, K., Moore, B. C. J., and Stone, M. A. (**2008**). "Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," J. Acoust. Soc. Am. **123**, 1140–1153.

Laneau, J., Moonen, M., and Wouters, J. (**2006**). "Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants," J. Acoust. Soc. Am. **119**, 491–506.

Langhans, A., and Kohlrausch, A. (**1992**). "Differences in auditory performance between monaural and diotic conditions. I: Masked thresholds in frozen noise," J. Acoust. Soc. Am. **91**, 3456–3470.

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

Lorenzi, C., Debruille, L., Garnier, S., Fleuriot, P., and Moore, B. C. J. (**2009**). "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal," J. Acoust. Soc. Am. **125**, 27–30.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (**2006**). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," Proc. Natl. Acad. Sci. U.S.A. **103**, 18866–18869.

McFadden, D. (**1986**). "Comodulation masking release: Effects of varying the level, duration, and time delay of the cue band," J. Acoust. Soc. Am. **80**, 1658–1667.

Moore, B. (**2008**). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," J. Assoc. Res. Otolaryngol. **9**, 399–406.

Nelson, P. B., Jin, S.-H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.

Qin, M. K., and Oxenham, A. J. (**2003**). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," J. Acoust. Soc. Am. **114**, 446–454.

Recio-Spinoso, A., Temchin, A. N., van Dijk, P., Fan, Y.-H., and Ruggero, M. A. (**2005**). "Wiener-Kernel analysis of responses to noise of chinchilla auditory-nerve fibers," J Neurophysiol **93**, 3615–3634.

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

Rosen, S. (**1992**). "Temporal information in speech: Acoustic, auditory and linguistic aspects," Philos. Trans. R. Soc. London, B, Biol. Sci. **336**, 367–373.

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.

Schooneveldt, G. P., and Moore, B. C. J. (**1989**). "Comodulation masking release (CMR) as a function of masker bandwidth, modulator bandwidth, and signal duration," J. Acoust. Soc. Am. **85**, 273–281.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.

Stickney, G. S., Zeng, F.-G., Litovsky, R., and Assmann, P. (**2004**). "Cochlear implant speech recognition with speech maskers," J. Acoust. Soc. Am. **116**, 1081–1091.

Verhey, J. L., Dau, T., and Kollmeier, B. (**1999**). "Within-channel cues in comodulation masking release (CMR): Experiments and model predictions using a modulation-filterbank model," J. Acoust. Soc. Am. **106**, 2733–2745.

Verhey, J. L., Pressnitzer, D., and Winter, I. M. (**2003**). "The psychophysics and physiology of comodulation masking release," Exp. Brain Res. **153**, 405–417.

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., and Cao, K. (**2005**). "Speech recognition with amplitude and frequency modulations," Proc. Natl. Acad. Sci. U.S.A. **102**, 2293–2298.

Zwicker, E. (**1952**). "Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones" ("The limits of audibility for amplitude modulation and frequency modulation of a pure tone"), Acustica, Akust. Beihefte **3**, AB125–AB133.