Forward masking of amplitude modulation: Basic characteristics[a]

Magdalena Wojtczak[b] and Neal F. Viemeister
Department of Psychology, University of Minnesota, 75 East River Rd., Minneapolis, Minnesota 55455

(Received 20 April 2005; revised 27 July 2005; accepted 1 August 2005)

In this study we demonstrate an effect for amplitude modulation (AM) that is analogous to forward making of audio frequencies, i.e., the modulation threshold for detection of AM (signal) is raised by preceding AM (masker). In the study we focused on the basic characteristics of the forward-masking effect. Functions representing recovery from AM forward masking measured with a 150-ms 40-Hz masker AM and a 50-ms signal AM of the same rate imposed on the same broadband-noise carrier, showed an exponential decay of forward masking with increasing delay from masker offset. Thresholds remained elevated by more than 2 dB over an interval of at least 150 ms following the masker. Masked-threshold patterns, measured with a fixed signal rate (20, 40, and 80 Hz) and a variable masker rate, showed tuning of the AM forward-masking effect. The tuning was approximately constant across signal modulation rates used and consistent with the idea of modulation-rate selective channels. Combining two equally effective forward maskers of different frequencies did not lead to an increase in forward masking relative to that produced by either component alone. Overall, the results are consistent with modulation-rate selective neural channels that adapt and recover from the adaptation relatively quickly. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.2042970]

PACS number(s): 43.66.Mk, 43.66.Dc [GDK] Pages: 3198–3210

I. INTRODUCTION

Envelope fluctuations provide important cues in the perception of auditory stimuli. Envelopes that differ in their temporal structure may help perceptually segregate sounds with overlapping spectra (Grimault *et al.*, 2002; Roberts *et al.*, 2002), and thus extract the target sound from the background. On the other hand, interference between concurrent temporal envelope fluctuations may adversely affect speech intelligibility, especially in listeners with reduced spectral resolution, such as hearing-impaired listeners (Hedrick and Jesteadt, 1996) or cochlear-implant users (Hedrick and Carney, 1997; Kwon and Turner, 2001).

Most sounds have envelopes that are complex waveforms and can be represented by a combination of sinusoidal amplitude modulations. An important question is whether the auditory system performs a spectral analysis on envelopes by processing different modulation components in modulation-rate selective channels or whether processing of envelopes is based on a purely temporal code.

Studies of modulation masking have demonstrated that the auditory system exhibits modulation-rate selectivity (Houtgast, 1989; Bacon and Grantham, 1989), albeit limited, as indicated by relatively broad masking patterns. This finding has inspired models of auditory processing that implement a bank of modulation filters at the output of individual peripheral channels (Dau *et al.*, 1997a, b; Ewert and Dau, 2000). Modulation-rate selective channels could be realized by neurons that have bandpass modulation transfer characteristics. Such neurons have been found in the cochlear nucleus (Frisina *et al.*, 1990; Rhode and Greenberg, 1994), in the inferior colliculus (Rees and Møller, 1983; Langner and Schreiner, 1988), and in the auditory cortex (Schulze and Langner, 1999; Liang *et al.*, 2002). Langner and Schreiner (1988) reported systematic mapping of best modulation frequencies (BMFs) in the inferior colliculus (IC) of the cat. That mapping, however, is not robust to changes in carrier level. Krishna and Semple (2000) observed that the BMF in the IC varies substantially across carrier levels, even over a relatively small (20-dB) range. Interpretation of the physiological data is further complicated by the fact that neural representations at all processing stages above the auditory nerve are nonhomogeneous and that different cell types exhibit different characteristics in their response to AM stimuli (for a review see Joris *et al.*, 2004). It is also unclear which characteristics of the response to AM are important for the perception of envelope fluctuations. Comparisons between psychophysical data from a variety of tasks involving modulation detection and discrimination and different response characteristics measured physiologically are, therefore, crucial for elucidating the mechanisms that underlie AM processing.

In this study we demonstrate a forward-masking effect in which a highly detectable AM raises the threshold for the detection of subsequently presented AM. The effect is analogous to forward masking in the domain of audio frequencies. Nonsimultaneous interference in modulation detection has been observed in numerous studies of adaptation to AM (Kay and Matthews, 1972; Regan and Tansley, 1979; Tansley and Suffield, 1983). Experiments in those studies were designed

---

[a] Portions of these data were presented at the XXVII Midwinter Research Meeting of the ARO (M. Wojtczak and N. F. Viemeister, Abstract No. 838).

[b] Send correspondence to Magdalena Wojtczak, Psychoacoustics Lab, Elliott Hall, N218, 75 East River Rd., Minneapolis, Minnesota 55455.

to produce the maximum amount of adaptation and thus, very long durations of adapting modulations were used (20–30 min). It was found that such long exposures to a sinusoidal AM decreased sensitivity to an ensuing modulation of the same rate. It is not clear whether the observed effect of adaptation plays any role in the perceptual processing of envelopes, since only negligible threshold elevation was observed after just a few minutes of exposure (Tansley and Suffield, 1983). More recently, Sheft (2000) observed significant threshold elevation (up to 7 dB) using a 3-min adapting AM before trials began and then interspersing the adaptor for 16 s between trials. However, long durations of uninterrupted exposure to a constant pattern of a temporal envelope are not commonly encountered. In addition, it has been questioned whether or not the adaptation paradigm is valid for elucidating the mechanisms responsible for encoding and processing AM. It has been demonstrated that for FM and frequency sweeps, the loss of sensitivity due to long prior exposure to these features, considered initially as reflecting neural adaptation, disappears with training (Moody *et al.*, 1984), or is small enough to be explained in terms of nonsensory factors (Wakefield and Viemeister, 1984). Recently, Lorenzi *et al.* (2004) measured adaptation to AM and found that a post-exposure impairment of AM detection, although initially in agreement with the data reported in earlier studies of adaptation, also disappears after extensive training.

Early studies that measured adaptation to AM have used the same modulation rate for the adapting AM and the signal to be detected. Thus, modulation-rate selectivity of the adaptation effect was not investigated. There is a hint in the study by Richards *et al.* (1997) that the effect of adaptation in AM detection is modulation-rate selective. They used only two modulation rates of the adapting AM: one that was equal to the modulation rate of the test AM (16 Hz), and the other that was substantially different from the test AM (56 Hz), and found that only the 16-Hz adaptor significantly decreased sensitivity to the 16-Hz test AM. However, their data do not allow for an inference about the sharpness of tuning. More recently, rate selectivity of adaptation to AM was shown by Sheft (2000) for AM detection (thresholds were elevated when the rate separation between the adapting and test AM was within one to two octaves) and by Wojtczak and Viemeister (2003) for the effect of an adapting AM on the perception of *suprathreshold* modulation depth. The tuning they observed was sharper than that observed in simultaneous modulation masking.

Forward masking of AM newly demonstrated in this study may be potentially very relevant to the processing of real-world auditory stimuli since the effect can be observed for relatively short durations of masking modulations (e.g., 150 ms), durations comparable to those observed in speech sounds. Our main focus in this study was to determine the basic characteristics of the effect, primarily the temporal and spectral region over which prior exposure to AM affects sensitivity to ensuing AM.

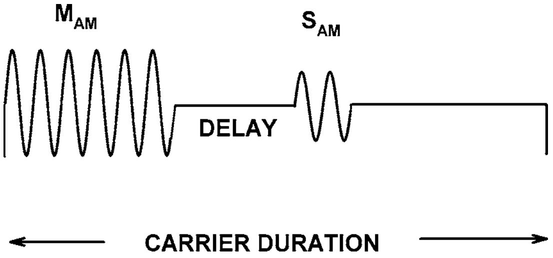

FIG. 1. Schematic illustration of the stimulus envelope in the signal interval. The envelope contains the masker AM (M_{AM}) and signal AM (S_{AM}) imposed on the same uninterrupted carrier.

II. EXPERIMENT 1: RECOVERY FROM AM FORWARD MASKING

Our purpose in this experiment was to measure the shape of the recovery function for AM forward masking. Figure 1 schematically illustrates the envelope of a stimulus containing both the masker AM (M_{AM}) and signal AM (S_{AM}). The detection of S_{AM} was measured as a function of the delay between the offset of M_{AM} and the onset of S_{AM}. To ensure that strong envelope fluctuations associated with gating the carrier off and on did not affect detection of the signal AM, both the masker and signal AM were sequentially imposed on the same uninterrupted carrier.

When attempting to measure the recovery function, it is desirable to use a very short signal so that the function can be measured with good resolution. This requires using a high modulation rate, since at least one cycle of modulation needs to be presented. However, the detection of modulation deteriorates with a decreasing number of modulation cycles (Sheft and Yost, 1990) and, for noise carriers, it also deteriorates with an increasing modulation rate above about 30 Hz (Viemeister, 1979). Thus, using just one cycle of a very high modulation rate could impose an artificial limit on the observed size of AM forward masking because it would limit the available range of modulation depths above the unmasked AM detection threshold. For these reasons, very low and very high modulation rates could not be used in the experiment.

A. Stimuli and procedure

The detection of a 40-Hz sinusoidal AM (S_{AM}) was measured using an adaptive three-interval forced-choice (3IFC) procedure that estimated the 79.4% point on the psychometric function (Levitt, 1971). Feedback indicating the correct response was provided on each trial. The signal modulation S_{AM} was presented alone in one condition, and after a 40-Hz M_{AM} in another condition. In the signal interval, the modulations were imposed sequentially on an uninterrupted noise carrier whose bandwidth extended from 0.1 to 10 kHz. The noise carrier had a spectrum level of 25 dB SPL measured at 1 kHz. For delays between S_{AM} and M_{AM} up to 210 ms, a carrier duration of 500-ms was used. A longer duration of the carrier (750 ms) had to be used when it was observed that the AM masked threshold was still elevated when S_{AM} was separated from M_{AM} by the longest delay permissible with the 500-ms carrier duration. The stimulus envelope shown in Fig. 1 was described by

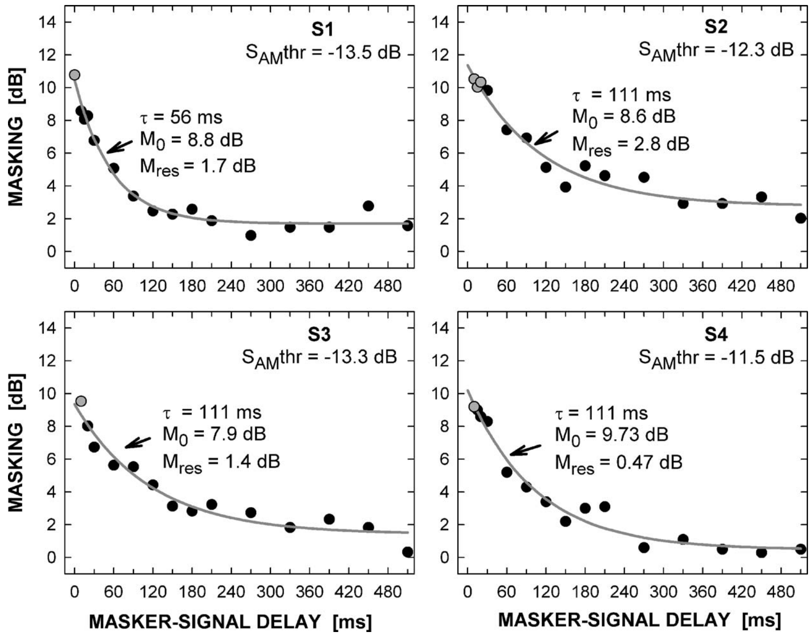

FIG. 2. AM forward masking (symbols) measured as a function of the delay between the offset of M_{AM} and the onset of S_{AM}. Gray lines represent the best fitting function described by Eq. (2). The amount of masking above the residual at the delay of 0 ms, the residual amount of masking, and the time constant describing recovery are provided in each panel. Also, the threshold for detecting unmasked S_{AM} is shown for each subject.

$$e(t) = 1 + m \sin(2\pi f_m t + \Phi) \text{ during the modulated portions of the carrier, and}$$

$$e(t) = 1 \text{ during the unmodulated portions,} \qquad (1)$$

where m, f_m, and Φ are the modulation depth, modulation rate, and the modulation starting phase, respectively. M_{AM} occupied the initial 150 ms of the carrier and its modulation depth was 1 ($20 \log m = 0$ dB). S_{AM} had a duration of 50 ms and its modulation depth was varied adaptively to find the threshold. Both M_{AM} and S_{AM} started at phase $\Phi = 0$ rad. Thresholds were measured for delays between the offset of M_{AM} and the onset of S_{AM}, ranging from 0 to 510 ms. When M_{AM} was not present, the unmasked detection of S_{AM} was measured with S_{AM} starting 150 ms after the carrier onset. The adaptive tracking procedure used a 2-dB step ($20 \log m$) until the first four reversals were obtained. The step size was reduced to 1 dB afterward. A total of 12 reversals were obtained and the threshold estimate was computed as the mean of the last eight reversals. Six to nine threshold estimates were averaged to compute the final threshold for each masker-signal delay. The run was aborted when the adaptive procedure called for a modulation depth greater than 0 dB, to avoid overmodulation. For delays at which more than three runs were aborted, thresholds were deemed immeasurable. When three or fewer runs were aborted, the threshold was computed as the mean of the remaining three to six runs. Those thresholds are distinguished by using gray symbols.

For delays up to 210 ms, the forward-masking functions were also measured with two types of a cue that was used to help perceptually segregate S_{AM} from M_{AM}. This was done to evaluate a potential role of temporal confusion that has been shown to affect the amount of forward masking in the audio-frequency domain (Moore and Glasberg, 1982; Neff, 1985, 1986). In one case a 4-kHz tone was presented ipsilaterally at 75 dB SPL and gated (with 5-ms ramps) for the duration of M_{AM}. In another case, an independent sample of the noise carrier modulated by M_{AM} was presented to the contralateral ear in each observation interval.

The stimuli were generated digitally on a PC using a 24-bit soundcard (Echo-Gina 24/96) and a sampling rate of 44.1 kHz. The noise carrier was generated in the frequency domain, with the components outside the passband set to zero. The amplitude and phase of the components inside the passband were randomly selected from the Rayleigh and rectangular distributions, respectively. The real part of the inverse FFT was then multiplied by the modulating waveforms. A different sample of noise was presented in each observation interval. The stimuli were presented monaurally to the left ear via Sony MDR-V6 earphones, except the condition where the noise carrier modulated by M_{AM} was presented to the opposite ear to eliminate possible temporal confusion.

B. Subjects

Six listeners with normal hearing participated in experiment 1. Four listeners were used to measure the recovery functions and two additional listeners were used to compare recovery for delays up to 210 ms with versus without a cuing stimulus. The listeners had audiometric thresholds to within 10 dB of laboratory norms at octave frequencies between 250 and 8000 Hz (thresholds better than 15 dB HL). Subject S3 (the first author) was highly trained in psychoacoustic tasks. The other subjects were naïve and were given at least 10 h training on selected conditions before data collection commenced.

C. Results

Modulation thresholds for detecting the unmasked 40-Hz S_{AM} were subtracted from thresholds measured for S_{AM} presented after M_{AM} to compute the amount of AM forward masking. Figure 2 shows the amount of AM forward

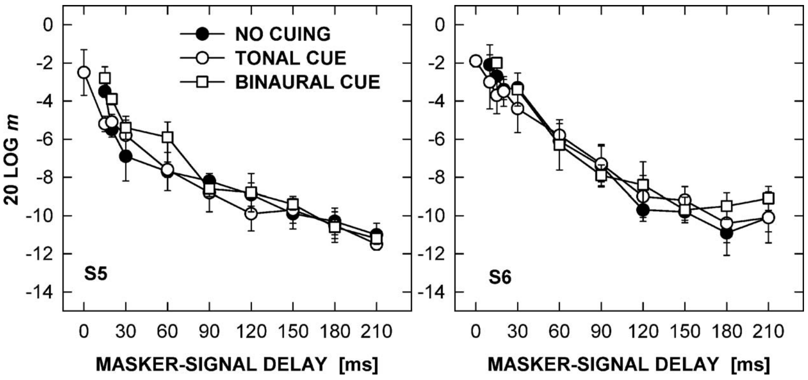

FIG. 3. Thresholds for the detection of S_{AM} plotted as a function of the delay between the offset of M_{AM} and the onset of S_{AM} measured without a cuing stimulus (filled circles), with a 4-kHz tone gated for the duration of M_{AM} (open circles), and with a contralateral noise modulated by M_{AM} (open squares). The error bars in this and all subsequent figures indicate one standard deviation. Data from two subjects.

masking (symbols) as a function of the delay between the offset of M_{AM} and the onset of S_{AM}. Standard deviations for the masked thresholds (not shown) estimated from six to nine runs were less than 1 dB in 94% of all cases (all delays and subjects). The maximum standard deviation did not exceed 1.2 dB. The unmasked thresholds are shown in the insert of each panel (S_{AM} thr). The mean unmasked threshold of −12.7 dB obtained by averaging thresholds from the four listeners is about 2 dB higher than the mean threshold for two listeners measured by Sheft and Yost (1990) using two cycles of 40-Hz AM imposed on a noise carrier. The difference could reflect the effect of a longer fringe (500 ms versus 150 ms used in this study) or slightly worse average sensitivity of our listeners.

For the 0-ms delay, a valid estimate of the masked threshold could not be obtained from any of the subjects. Subject S1 was the only one who successfully completed three out of six runs, but three other runs were aborted due to the adaptive procedure calling for a modulation depth greater than 0 dB. Since the duration of S_{AM} was 50 ms, this result indicates that over a period of 50 ms listeners could not reliably detect a 40-Hz AM when it followed AM of the same rate. Similarly, some runs were aborted at other short delays for subjects S2 (for a 10, 15, and 20-ms delay), and S3, S4 (for a 10-ms delay), indicated by gray symbols. Generally, the masking effect decays rapidly within the first 100–150 ms (offset to onset) and then gradually asymptotes. Subjects S1 and S2 exhibited a small residual amount of masking at the longest delay (510 ms).

To determine the time constant characterizing recovery from AM forward masking, the data in Fig. 2 were fitted with an exponential function given by the following equation:

$$M(d) = M_0 \cdot e^{-d/\tau} + M_{\mathrm{res}}, \tag{2}$$

where $M(d)$ is the amount of masking in dB at a delay d, M_{res} is the residual masking, M_0 is the amount of masking at a 0-ms delay relative to the residual masking, and τ is the time constant describing the rate of recovery. The parameters yielding the best fit are given in the insert of each panel. The proportion of variance accounted for by the exponential fit (r^2) was 0.98 for S1, 0.96 for S2, 0.95 for S3, and 0.97 for S4.

Since the rates of M_{AM} and S_{AM} were identical, the masked thresholds measured at the shortest delays might have been artificially raised by the listeners' inability to determine when the masker ended and the signal began. Thresholds affected by temporal confusion are expected to decrease when a stimulus cuing the temporal end of the masker is used (Moore and Glasberg, 1982). Figure 3 shows recovery functions measured in two subjects with two types of cuing stimuli added to investigate a potential role of temporal confusion. Neither subject could detect the signal without a cuing stimulus for the 0-ms delay and both showed improvement in this condition when the 4-kHz cue was presented during the masker modulation. Subject S6, however, still required a modulation depth greater than 0 dB in two out of six runs. S5 also showed a slight (less than 2 dB) decrease in threshold for a 15-ms delay, and S6 showed a less than 1-dB improvement for 10- and 15-ms delays when the cuing tone was present. The cuing tone led to an improvement in performance in no other condition. The contralateral noise (open squares) did not reduce the threshold at any delay. For shorter delays, it may have provided distraction, causing a slight increase in forward masking compared with that observed without the noise. Overall, the data suggest that the observed AM forward masking is not affected by temporal confusion, at least for delays greater than 15 ms. It is possible, however, that the cues were not effective enough in eliminating temporal confusion (this issue is addressed in Sec. IV).

After data collection was completed, at least 30 additional blocks were run in selected conditions to make sure that the amount of masking could not be eliminated or substantially reduced by training. The data showed that the training given prior to data collection was sufficient and the amounts of masking presented in Fig. 2 likely reflect the asymptotic performance.

III. EXPERIMENT 2: RATE SELECTIVITY IN AM FORWARD MASKING

Masking patterns obtained for the simultaneously presented masker and signal AM show a clear peak for signal modulation rates that are similar to the masker modulation rates (Bacon and Grantham, 1989; Houtgast, 1989; Takahashi and Bacon, 1992; Dau *et al.*, 1997a, b; Ewert and Dau, 2000; Ewert *et al.*, 2002). This tuning of masking may reflect the processing of AM by rate-selective neural channels. However, some aspects of modulation masking cannot be explained by spectral models of masking. For example, when the rate of the masker AM is low and the rate of the signal AM is much higher, the presence of the masker facilitates signal detection and thus leads to thresholds that are lower than the unmasked thresholds. The "negative masking," seen in the data of Bacon and Grantham (1989) and Strickland and Viemeister (1996), was explained in terms of detecting "local temporal features." This strategy for detecting the signal would require that listeners are able to evaluate their decision variable over short temporal intervals ("looks") and improve their performance by detecting the signal in the troughs of the masker modulation, where the effective modulation depth of the signal is the greatest. The same explanation could account for the fact that for a masker rate much lower than the signal rate, detection improves as the modulation depth of the masker increases, which is also seen in the two studies. Another aspect of AM masking that cannot be explained in terms of spectral models is that simultaneous AM masking depends strongly on the relative starting phases of the signal AM and the masker AM when the signal and masker rates differ by a factor of 2 (Strickland and Viemeister, 1996). Moreover, in simultaneous AM masking, listeners can improve their performance by detecting beats between the masker and signal AM component (Strickland and Viemeister, 1996; Ewert *et al.*, 2002). The beat rate, although not represented in the modulation spectrum, can be introduced in the internal representation of the envelope as a result of compressive nonlinearity and possibly other nonlinearities in the auditory system (Shofner *et al.*, 1996; Moore *et al.*, 1999; Füllgrabe *et al.*, 2005).

Using a forward-masking paradigm to measure tuning prevents listeners from detecting the signal AM in local dips of the masker AM and from using any other local temporal cues, because the two modulations do not overlap in time. In addition, modulation distortion products are unavailable as cues when the masker and signal AM are temporally separated. Thus, forward masking may potentially provide a valid assessment of tuning in AM processing, if it is assumed that the mechanism underlying the forward masking does not depend on the modulation rate of the masker (an assumption that has not been verified for AM).

In this experiment, the detection of S_{AM} was measured as a function of the modulation rate of M_{AM} for a short (fixed) temporal delay between S_{AM} and M_{AM}.

A. Stimuli and procedure

Masked-threshold patterns (with the signal rate fixed and the masker rate varied) were measured for S_{AM} rates of 20, 40, and 80 Hz. In all cases, S_{AM} had a duration of 50 ms so the number of cycles in the signal increased from 1 to 4 as the modulation rate increased. The noise carrier had the same bandwidth as that used in experiment 1, and was gated for 600 ms. The masker AM had a duration of 500 ms and a modulation depth of 0 dB (100% modulation). A masker duration of 500 ms was necessary so that at least one cycle of M_{AM} could be presented for each masker modulation rate. The detection of S_{AM} was measured for modulation rates of M_{AM} selected from the range between 2 and 256 Hz in one-octave steps. The delay between the offset of M_{AM} and the onset of S_{AM} was fixed for a given S_{AM} rate but was not always the same across subjects or S_{AM} rates. In each case, the shortest delay was used for which a subject was able to perform the task for the M_{AM} rate that produced most masking (which happened sometimes for unequal rates of M_{AM} and S_{AM}). All delays were in the range between 20 and 40 ms.

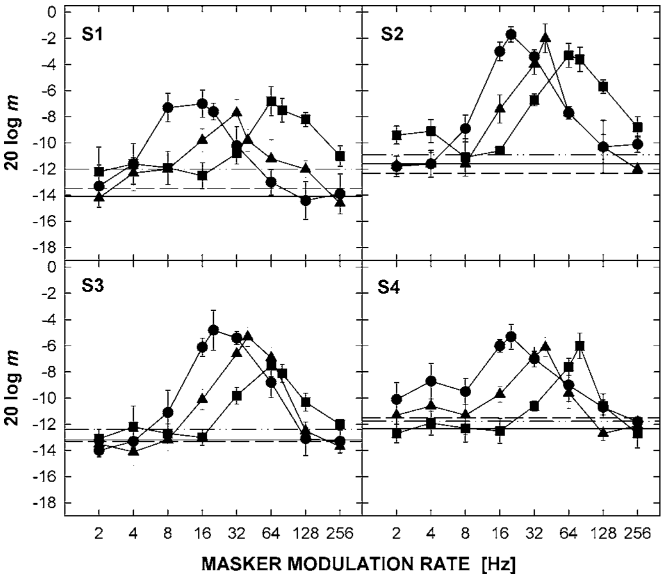

FIG. 4. Thresholds for the detection of S_{AM} plotted as a function of the rate of M_{AM}, for S_{AM} rates of 20-Hz (circles), 40 Hz (triangles), and 80 Hz (squares). Also shown are unmasked thresholds for the S_{AM} of 20 Hz (solid line), 40 Hz (dashed line), and 80 Hz (dotted-dashed line). In most cases the delay between M_{AM} and S_{AM} was fixed at 20 ms, except the 20-Hz data were collected at a 30-ms delay for S2 and S4, and the 80-Hz data were collected at a 40-ms delay for S2, and a 30-ms delay for S4.

The experimental procedure, generation, and presentation of the stimuli were the same as in experiment 1.

B. Subjects

The four subjects for whom recovery functions were measured in experiment 1 participated in this experiment.

C. Results

Figure 4 shows forward-masked AM thresholds, for all three modulation rates of S_{AM}, plotted as a function of the modulation rate of M_{AM}. The horizontal lines indicate the unmasked thresholds. The data reveal tuning of the forward-masking effect. The masker-signal delays at which the masked-threshold patterns were measured are given in the caption of Fig. 4. In some cases, the peak of the masking function is shifted toward lower rates relative to the rate of S_{AM} (S1 for the 20- and 80-Hz S_{AM}, and S2 and S3 for the 80-Hz S_{AM}). The peaks of the masking patterns are generally similar across the three modulation rates of S_{AM}, with the exception of S2 and S3, who showed a slightly lower peak threshold for the 80-Hz rate. There is a slight trend for the amount of masking, computed as a difference between the masked and unmasked threshold, to decrease with the increasing modulation rate of S_{AM}.

Not surprisingly, negative masking is not observed when the masker rate is much lower than the signal rate. Since the masker and the signal were not presented simultaneously,

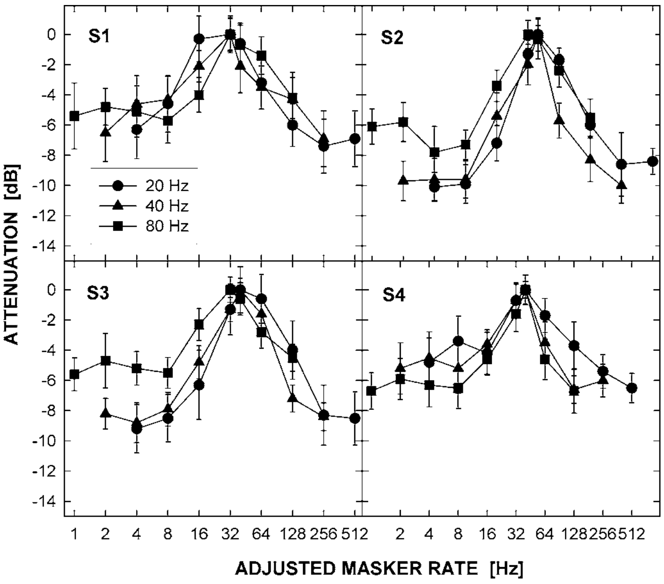

FIG. 5. Attenuation functions derived from data in Fig. 4. Functions for the 20-and 80-Hz S_{AM} are shifted to have their peak correspond to 40 Hz for the purpose of comparison.

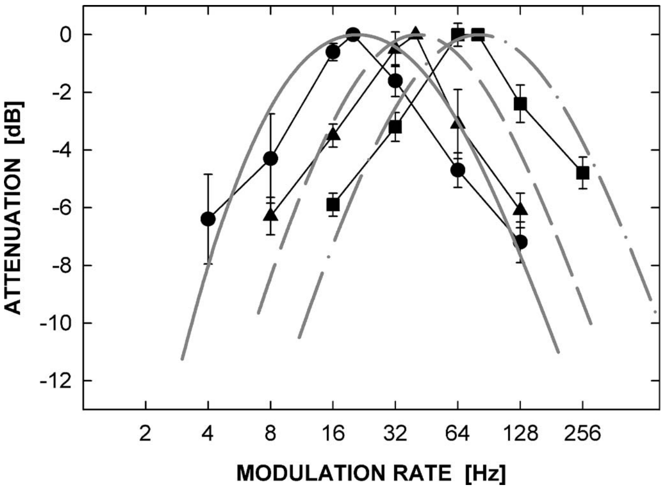

FIG. 6. Attenuation functions averaged across listeners for S_{AM} rates of 20, 40, and 80 Hz. The gray lines show the shapes of second-order bandpass filters fit to the data.

listeners could not take advantage of the larger local effective modulation depth of S_{AM} in the troughs of M_{AM}. No other local temporal features could be used to improve performance for M_{AM} differing in rate from S_{AM} relative to the performance observed for equal rates. Thus, the tuning of the forward-masking effect suggests that envelope processing is performed by rate-selective channels.

Figure 5 shows attenuation functions derived from the masked-threshold patterns. The functions were derived assuming that a masked threshold is obtained when the ratio of the S_{AM} amplitude to the M_{AM} amplitude at the output of the channel processing the signal reaches a constant criterion value. It was also assumed that the masker that produces the most masking is not attenuated by the channel processing the signal (0-dB attenuation). As mentioned above, the maximally effective masker was not always the on-frequency masker. Under these assumptions, the attenuation functions approximate the shape of hypothetical modulation filters, for the three modulation rates used in this study. To facilitate a comparison of the sharpness of tuning at the three modulation rates, the attenuation functions shown in Fig. 5, for the 20- and 80-Hz rates, are shifted upward and downward, respectively, to have their peak at the 40-Hz rate. The three functions would overlap if the filters had a constant Q value. Since the tails of the functions are limited by the unmasked threshold, the sharpness of tuning should be compared only around the peak, i.e., for the adjusted masker rates between 8 and 128 Hz. The data in Fig. 5 indicate that there is no systematic dependence of the relative bandwidth of a hypothetical modulation filter on the center frequency of the filter, over the range of the rates tested.

To quantify tuning, a second-order bandpass Butterworth filter, symmetric on the log frequency scale, was fitted to the adjusted attenuation functions in the range between 8 and 128 Hz. MATLAB's *fminbnd* function was used to find a ratio of the filters' cutoff frequencies that produced the smallest rms error between the data and the fitted filter. The Q values of the filters fitted to the individual data and the rms errors of the fits are given in Table I.

Figure 6 shows attenuation functions obtained by averaging the amounts of attenuation across the four listeners separately for each of the three signal modulation rates. The gray lines show the attenuation characteristics of the second-order Butterworth filters fitted to the averaged data. The fitted filters are clearly too broad near the peak, which results from fixing the order of the fitted filters at 2. This was done in order to compare the Q values of the fitted filters with those used to fit AM masking data in studies that measured simultaneous masking (Ewert and Dau, 2000). A higher order of the filter would produce a better fit to the data. As in studies of AM simultaneous masking, the filters were as-

TABLE I. Q values of second-order bandpass filters and the rms errors of the fits to the individual and mean data for S_{AM} rates of 20, 40, and 80 Hz.

	20 Hz		40 Hz		80 Hz	
	Q	rms_err	Q	rms_err	Q	rms_err
S1	0.35	1.45	0.36	1.40	0.41	1.12
S2	0.62	1.30	0.71	2.12	0.50	0.71
S3	0.47	1.31	0.55	1.28	0.41	0.76
S4	0.35	2.10	0.42	2.02	0.53	2.06
Mean	0.36	1.21	0.44	1.56	0.41	1.06

TABLE II. Thresholds and standard deviations for detecting S_{AM} for different starting phases.

	$-\pi/2$	0	$\pi/2$	π
S1	−6.6 (0.8)	−6.1 (0.8)	−6.2 (0.4)	−6.0 (0.4)
S2	−2.5 (0.7)	−2.2 (1.3)	−2.3 (0.9)	−2.1 (1.2)
S3	−6.6 (0.9)	−5.9 (0.6)	−6.2 (1.0)	−6.0 (0.5)

sumed to be symmetrical in log frequency. This assumption also contributed to the fitting error, since the measured attenuation functions for the 40- and 80-Hz S_{AM} appear to have a slightly steeper slope on the high-frequency side than on the low-frequency side, when plotted against log frequency. The Q values and the rms errors for the filters fitted to the averaged functions are shown in the bottom row of Table I. The Q values indicate very broad tuning. Data from tasks measuring simultaneous masking of AM suggest Q values between 1 and 2 (Dau *et al.*, 1997a; Ewert *et al.*, 2002). The much sharper tuning in simultaneous masking may result from listeners' ability to use local temporal features or distortion modulation components to enhance their performance in cases where M_{AM} and S_{AM} have different rates.

Caution is necessary in speculating about the mechanism underlying forward masking of AM, since the exact mechanism involved in coding AM in the auditory system has not been yet elucidated. One possibility is that it reflects ringing of the channels processing AM. Although broad tuning observed in modulation masking implies very short ringing times, a control condition was run to evaluate a potential effect of the modulation starting phase on the detection of S_{AM} (Sek and Moore 2002). Thresholds were measured at a fixed delay between M_{AM} and S_{AM} of 30 ms, with the S_{AM} starting phase set to $-\pi/2$, 0, $\pi/2$, and π rad. The 0-rad condition (the data point for the 30-ms delay in experiment 1) was rerun to assess a potential training effect. Data from three listeners (S4 was unavailable for testing) are shown in Table II. Although thresholds appear consistently lower for $-\pi/2$ and $\pi/2$ than for the 0 and π phases, a one-way ANOVA showed that the effect of the starting phase was not significant ($F=0.08, p=0.97$). This result suggests that the forward-masking effect does not result from a simultaneous interaction between the masker and signal AM due to ringing, in which information about the modulation phase is preserved.

IV. EXPERIMENT 3: MASKING BY COMBINED AM FORWARD MASKERS

This experiment compares the amount of masking produced by an on-frequency masker (40-Hz M_{AM}) with the amount of masking produced by that same masker paired with a masker having a different modulation rate. The aims of this experiment are (1) to test the hypothesis that the size of the forward-masking effect is determined by the total power of the modulation masker at the output of the channel tuned to the signal modulation rate; (2) to further test to what extent the perceptual similarity between M_{AM} and S_{AM} affects tuning observed in AM forward masking.

Regarding the first aim, results from previous experiments suggest that modulations with different rates are processed in modulation-rate selective channels. In this experiment, additivity of forward masking is tested by comparing the amounts of masking produced by simultaneously present similarly effective maskers with the amount of masking produced by either masker presented separately. If the size of forward masking is determined by the total power of the modulation masker at the output of the channel tuned to the signal modulation rate, then combining two maskers that separately produce similar amounts of masking should result in an increased masked threshold compared with that produced by either masker alone.

Regarding the second aim, combining a 40-Hz AM with a distinctly different rate may facilitate the perceptual segregation between the masker and signal AM. In the second part of experiment 1, a 4-kHz tone was presented during M_{AM} to mark the temporal end of the masker. Such a cue was expected to reduce a potential effect of temporal confusion between M_{AM} and S_{AM}, when the two modulations are similar perceptually. It is possible that the tonal cue was not sufficient to help listeners distinguish between the masker and the signal. In that case, tuning of the masking effect could reflect

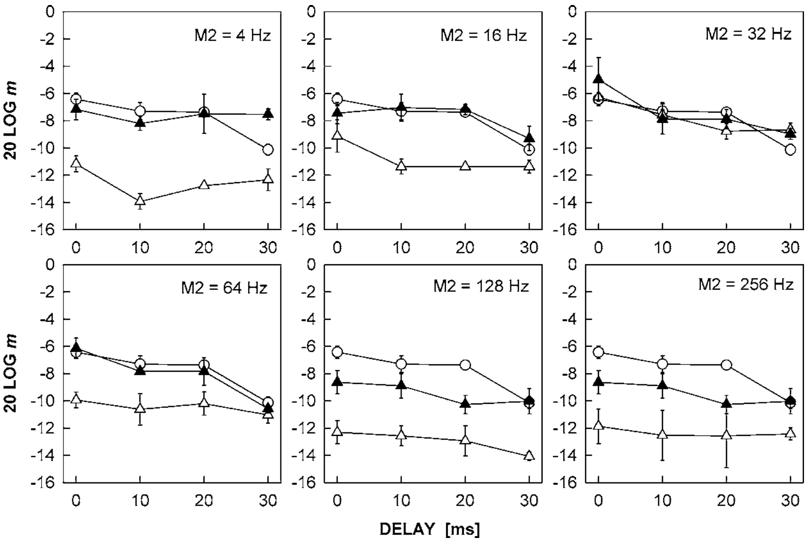

FIG. 7. Thresholds for detecting 40-Hz S_{AM} presented after a single-component M_{AM} (open symbols) and after a two-component M_{AM}. Data for the 40-Hz masker alone are shown by open circles, data for an off-frequency masker alone are shown by open triangles, and data for the two sinusoidal maskers combined are shown by filled triangles. Data for S1.

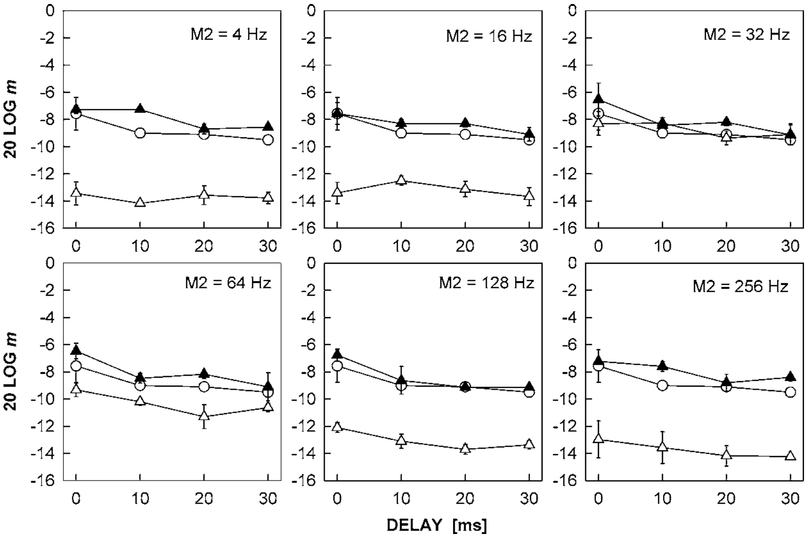

FIG. 8. The same as in Fig. 7, but for listener S3.

a decreasing perceptual similarity between M_{AM} and S_{AM} as the difference between their rates increases. Adding a modulation that has a different modulation rate from S_{AM} (one that does not produce any forward masking of S_{AM} on its own) to an on-frequency M_{AM} (with the same rate as that of S_{AM}) may provide a more efficient cue for perceptual segregation between M_{AM} and S_{AM}. If that were the case, and if forward masking reflected perceptual confusion, then the threshold would be expected to decrease in the presence of the modulation added to the on-frequency masker. Thus, measuring masking by two combined maskers could provide an insight into the nature of the forward-masking effect.

A. Stimuli and procedure

The experimental procedure, the carrier, and the durations of the masker and signal modulations were the same as in experiment 2. The masker modulation was either sinusoidal AM or a combination of two sinusoidal AMs. The combined masker AM consisted of a 40-Hz sinusoid paired with one of the following modulation rates: 4, 16, 32, 64, 128, and 256 Hz. In all conditions, each sinusoidal AM had a modulation depth of −6 dB ($20 \log m$). All signal and masker modulations started with a phase of 0 rad. Thresholds for detecting S_{AM} were measured for four delays between the masker and signal, 0, 10, 20, and 30 ms. The method for the generation and presentation of the stimuli was the same as in the previous experiments.

B. Subjects

Three listeners completed the task. All of them participated in the first two experiments. Subject S2 was not available for testing at the time of the experiment.

C. Results

Data from the three listeners are shown in Figs. 7–9. Open symbols show data for individual maskers presented alone and filled symbols show thresholds obtained for two maskers combined. For the 40-Hz masker, the amount of masking is less than that measured in experiment 1. This is not surprising because the modulation depth of M_{AM} used here was 6 dB lower than that used in the previous experiments. Since the 32-Hz AM produced an amount of masking most similar to that observed for the 40-Hz masker, the ad-

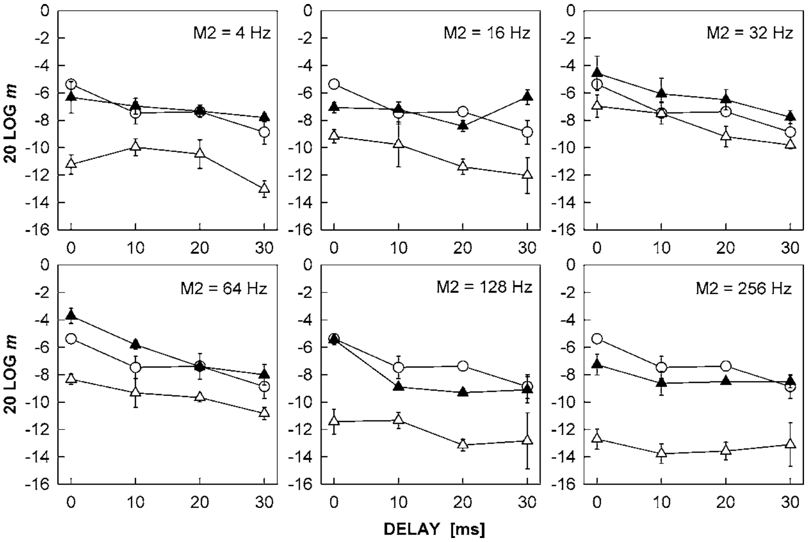

FIG. 9. The same as in Fig. 7, but for listener S4.

ditivity of masking would be expected to be revealed when these two masker rates are combined. A two-way ANOVA was performed separately on each listener's data with the masker modulation rate and the masker-signal delay as the main factors. For the 32- versus 40-Hz masker presented separately, the ANOVA revealed no significant effect of the masker rate for S1 ($F=0, p=0.96$) and S3 ($F=0.02, p=0.90$), but listener S4 showed significantly more masking produced by the 40-Hz rate ($F=17, p=0.001$). For equally effective maskers, a simple additivity of envelope power would predict a 3 dB increase in the masked threshold. The data do not show an increase of forward masking when the 32- and 40-Hz maskers were presented together, in comparison with the forward masking produced by the 40-Hz masker alone. A two-way ANOVA with the type of masker (40 Hz alone versus 32+40 Hz), and the delay as the main factors revealed no significant effect of the masker type for listener S1 ($F=0.47$, $p=0.50$) and S3 ($F=2.5$, $p=0.13$) and only a marginally significant increase in forward masking for S4 ($F=10.11$, $p=0.006$). There was no significant interaction between the delay and the masker type for any of the three listeners. The lack of a significant increase in masking produced by two equally effective maskers compared with either masker alone argues against the possibility that forward masking depends on the total power of the forward masker at the output of the modulation channel processing the signal rate.

For M_{AM} other than 32 and 40 Hz presented alone, masked thresholds were substantially lower than the thresholds for the 40-Hz masker presented alone, reflecting tuning in AM forward masking. Thus, pairing an off-frequency masker with the 40-Hz masker should result in masked thresholds that are similar to those produced by the more effective masker. This was generally the case, except for some conditions where combining the 40-Hz masker with maskers of higher rates (128 and 256 Hz) led to some improvement in performance relative to that for the 40-Hz masker alone. This effect was the strongest for listener S1 (Fig. 7) at the three shortest delays, 0, 10, and 20 ms, but S4 (Fig. 9) also showed a slight decrease in threshold for the same delays when the off-frequency rate was 256 Hz, and for the 10- and 20-ms delays when the off-frequency rate was 128 Hz. The possible reasons for this improvement in signal detection will be discussed in the following section. In contrast, S3 (Fig. 8) showed a slight elevation in threshold when the off-frequency masker was added, but the change in threshold was only marginally significant ($F=12.5$, $p=0.003$).

V. DISCUSSION

There is ample evidence demonstrating masking between concurrent sinusoidal modulations of similar rates (Houtgast, 1989; Bacon and Grantham, 1989; Dau *et al.*, 1997a, b; Ewert and Dau, 2000; Ewert *et al.*, 2002). This interference between spectral components within a temporal envelope determines the amount of temporal information that is available to the listener. In a dynamically changing envelope, that information might be further limited by interference between modulations that occur nonsimultaneously. Data presented in this study convincingly demonstrate an effect that is analogous to forward masking of audio frequencies. A highly detectable modulation raises the threshold for detection of 50-ms modulation with the same or a similar rate when the latter follows it immediately after delays less than about 150 ms. The effect is referred to here as AM forward masking because of the apparent analogy and because the exact mechanism underlying the effect is presently unknown (thus, the term "adaptation" has been avoided here).

A. The temporal extent of AM forward masking

Data from the first experiment show that listeners could not consistently detect a 50-ms 100% modulation that immediately followed a 150-ms modulation of the same rate. Thresholds decreased with an increasing delay between M_{AM} and S_{AM} but in all listeners, thresholds remained elevated by more than 2 dB over at least a 150-ms interval. It is unclear to what extent this forward-masking effect plays a role in the processing of real-world auditory stimuli. The function representing recovery from AM forward masking was measured for a modulation rate of 40 Hz and the tuning of the effect was measured for rates between 20 and 80 Hz. Although such high rates are not the most prevalent in speech, they have been shown to contribute to speech recognition, especially when all spectral cues are removed and listeners have to rely on information contained in the temporal envelope (van Tasell *et al.*, 1987). Durations of the masker and signal AM used in this study fall within the range of durations for speech tokens (e.g., vowels). For modulation rates that are dominant in speech (below 16 Hz), one cycle of modulation approaches the time interval over which the auditory system recovers from forward masking, and thus the AM forward masking may not appreciably affect their perception.

The recovery times estimated from an exponential fit to the data in experiment 1 were remarkably consistent across three listeners (111 ms). One subject showed a faster recovery (56 ms). The time constants characterizing recovery from AM forward masking in all four listeners are very similar to those describing recovery from forward masking of audiofrequencies for listeners with sensorineural hearing losses (Nelson and Freyman, 1987) and recovery from forward masking in electrical stimulation for cochlear-implant users (Nelson and Donaldson, 2002). Since the two groups of listeners have reduced or nonexistent peripheral compression, the time constants reported in those studies are not affected by the effect of peripheral compression on the slope of the forward-masking curve measured behaviorally (Plack and Oxenham, 1998). Thus, those time constants may accurately describe the recovery of the mechanism underlying forward masking. Under this assumption, it is tempting to speculate that the similarity between the time constants describing recovery from AM forward masking and those characterizing forward masking of audio frequencies in the auditory system without peripheral compression may suggest that the two effects have the same type of underlying mechanism, e.g., they both may be mediated by neural adaptation.

The relatively fast recovery from the forward-masking effect explains why Tansley and Suffield (1983) did not see a substantial change in threshold after a few minutes of uninterrupted exposure to AM, since they used a 7 s test AM. When the signal is substantially longer than the time that the system needs to recover from forward masking, no threshold elevation is observed because a large portion of the signal is unaffected by the masker. Thus, even if the amount of masking depends on masker duration, the signal was likely long enough to be detected at its unmasked threshold level.

B. Tuning of AM forward masking

When sinusoidal modulations with different rates are simultaneously present in an envelope, listeners apparently can use local temporal features to detect the signal modulation in the presence of a masker. There is evidence that listeners can take advantage of troughs in the masker modulation to detect higher-rate modulations that are effectively magnified during low-level portions of a fluctuating stimulus envelope (Strickland and Viemeister, 1996). The availability of such local temporal cues affects the shape of masked-threshold patterns, and thus the bandwidth of a modulation filter derived from the masking data. One might argue that a more accurate estimate of the filter shape and bandwidth could be obtained from forward-masking data. This would require an assumption that the rate of recovery does not depend on the frequency of the masker modulation. While this assumption has not been specifically tested, our limited data at masker-signal delays shorter than those used in experiment 2 suggest that it might be justified.[1] Functions relating the masked threshold to the modulation rate of the masker AM have a distinct peak around the rate of the signal AM. The second-order Butterworth filters fitted to the data in Fig. 6 reveal very broad tuning. The Q values of the filters fitted to the attenuation functions averaged across listeners were less than 0.5 at all three signal modulation rates. Although such low Q values have been reported in some studies that used noise carriers (Ewert and Dau, 2000), most data from simultaneous modulation-masking experiments imply sharper tuning of AM, described by Q values between 1 and 2 (Dau *et al.*, 1997a, b; Ewert *et al.*, 2002). The sharper tuning in simultaneous masking may reflect the availability of local temporal cues that may lead to an improvement in performance as the signal and masker rates become more dissimilar and availability of modulation distortion components that could lead to a sharper peak of the measured filter.

As shown in Fig. 6, the data that presumably approximate the modulation-filter shapes are slightly asymmetric on a log frequency scale, with a steeper high-frequency side. Such asymmetry was also revealed in the data of Ewert and Dau (2000). To model the asymmetry of attenuation functions obtained from their simultaneous-masking data, Ewert and Dau applied a low-pass filter with a cutoff frequency of 150 Hz to the output of a bandpass modulation filter that was symmetric on a log-frequency scale. The choice of the cutoff frequency was dictated by modulation-detection data obtained for high-frequency tonal carriers that suggest "sluggishness" in envelope processing for modulation rates above 150 Hz (Kohlrausch *et al.*, 2000). The attenuation functions measured in experiment 2 may reflect the combined symmetric bandpass processing of modulation rates and lowpass filtering due to the sluggishness of temporal processing, but to predict the asymmetry at the 40-Hz rate seen in Fig. 6, a cutoff frequency lower than 150 Hz would have to be used.

C. Ringing in modulation channels

An implication of processing AM by modulation-rate selective channels is that those channels have a finite impulse response (ringing) of the modulation channels. One reason for forward masking could be an overlap of M_{AM} and S_{AM} due to ringing in the filter processing both modulations. If the modulation phase was preserved in ringing and if ringing mediated the forward-masking effect, then the detection of S_{AM} would be expected to depend on its starting phase. Data shown in Table II demonstrate that was not the case; at a 30-ms delay, at which substantial forward masking was observed, threshold was independent of the S_{AM} starting phase. The interpretation of this result is not straightforward given that the mechanism for coding AM is still unknown. Physiological data suggest that the coding mechanism changes as the stimulus moves from peripheral processing stages to more central sites. At the auditory periphery, responses to AM exhibit a high degree of synchronization to the phase of the stimulating AM (Cooper *et al.*, 1993; Wang and Sachs, 1993), but at the level of the Inferior Colliculus (IC) that code becomes partially converted to one based on the average firing rate (Langner and Schreiner, 1988; Krishna and Semple, 2000). Subsequently, the rate-based code becomes dominant at the level of the Primary Auditory Cortex (e.g., Liang *et al.*, 2002). Although the synchrony-to-rate conversion in AM coding may not occur in some species [e.g., guinea pig, Middlebrooks (2005)], it is possible that such conversion occurs in the human auditory system. The tuning of rate-based responses may imply a more complicated form of ringing that would be insensitive to the phase of modulation. Thus, the lack of phase effects in forward masking does not allow for unequivocal rejection of a role of persistence of the response to M_{AM} beyond the physical offset of that modulation.

D. "Additivity" of AM forward masking

It is sometimes useful to use a simple model of channel processing such as the envelope power spectrum model (EPSM) proposed by Ewert and Dau (2000). When applied to simultaneous-masking data, the model does not produce accurate predictions when temporal cues (such as beats between modulation components or enhanced local modulation depth in the troughs of the masker AM) can be used to improve detection. Those temporal cues are not available in forward masking, and thus one might expect that a spectral model of AM processing could be more successful in accounting for forward-masking data.

The envelope power spectrum model assumes that the detection of a certain signal AM rate is determined by the total power of envelope components that are processed by the modulation filter tuned to the signal rate. Comparing the amount of forward masking produced by the 100% (0-dB) 40-Hz AM in experiment 1 with that produced by the 50% (−6 dB) AM of the same rate in experiment 3 suggests a decrease in the amount of forward masking with decreasing power of the masker. The EPSM would predict such a decrease. If the system were linear, then combining two equally effective maskers should produce a 3-dB increase in the masked threshold. This was not observed in experiment 3, where thresholds for the 32- and 40-Hz masker presented separately did not differ significantly, and, yet, combining the two maskers did not lead to an increase in the masked threshold. Thus, a model assuming the integration of envelope power within a modulation filter cannot predict the data from experiment 3.

Two listeners showed a decrease in the amount of forward masking when an off-frequency modulation producing very little or no masking on its own was added to the 40-Hz masker modulation. One explanation for this effect is in terms of temporal confusion. As mentioned in the introduction to experiment 1, forward masking at short delays may be elevated due to the listener's inability to temporally separate the masker from the signal. A pure tone presented during the masker modulation in experiment 1 produced only a slight improvement (1–2 dB) in performance at the shortest delays. It is likely that in the case of two modulation maskers combined, the high-frequency modulation also provides a cue by changing the percept of the masker AM and making it less similar to the signal AM. One might argue that modulations with very low rates should provide a similar cue that would lead to an improvement in signal detection, but this is not seen in the data in Figs. 7–9. In that case, however, listeners reported being able to hear the low-frequency modulation and the 40-Hz modulation as separate. This may result from the fact that the two modulations produce percepts that fall into separate categories, temporal following for the low rate, and roughness for the higher rate (Wright and Dai, 1998). Consequently, the low-frequency AM does not change the percept of the 40-Hz AM enough to eliminate confusion between the masker and the signal. In contrast, high off-frequency rates produce a percept that falls into the same category as that produced by the 40-Hz rate. Listeners reported that when the 128 or 256 Hz are combined with the 40-Hz AM, the 40-Hz modulation component could not be heard out separately, and thus the combined masker led to a percept that was dissimilar from the signal AM. The explanation based on a reduction of temporal confusion is supported by the fact that listener S3, whose performance for the 40-Hz masker alone was the best, did not show an improvement when a modulation that could help distinguish between the signal and masker AM was presented with the 40-Hz AM. For the other two listeners, no improvement was observed when the masker-signal delay was increased to 30 ms.

Another explanation is in terms of inhibition of the response to the 40-Hz AM by a higher-rate AM. Physiological measurements show that neurons at more central processing sites can respond strongly to a relatively narrow range of modulation rates, but the response of the same neurons shows strong inhibition when they are stimulated by other rates (Krishna and Semple, 2000). Although Krishna and Semple did not stimulate their neurons by two modulation rates simultaneously, it is plausible that in the case of simultaneous stimulation by a rate that produces a strong response and by an inhibitory rate, the resulting response would be reduced. Since the decrease in threshold for S1 and S2 depended on the masker-signal delay, the explanation based on inhibition is less likely.

E. Final remarks

The mechanisms underlying AM forward masking are uncertain, especially since the mechanisms involved in AM coding have not been determined. One possible explanation could be in terms of the adaptation of neurons responding to AM. A recent physiological study by Bartlett and Wang (2005) found an effect of prior exposure to AM on subsequently presented AM of the same and different rates in cortical responses of marmoset monkeys. Although AM durations they used were longer than those in the experiments conducted within this study (from 500 to 2500 ms), the data are consistent with the adaptation as a mechanism underlying forward masking in modulation processing. Bartlett and Wang found that for equal masker and signal rates, cortical responses to the ensuing AM were strongly suppressed, with the strength of that suppression decreasing as the temporal separation between the two modulations increased. They also observed the tuning of that effect. The apparent similarity between their and our data has to be treated with caution since the parameters of the stimuli were very different, and in their study the carrier was gated with each modulation, which was not the case in our experiments. No other physiological studies that would allow more elaborate speculations are available at present.

VI. SUMMARY

The observed results can be summarized as follows.

(1) Data from our experiments demonstrate an effect analogous to forward masking observed for auditory frequencies. The sensitivity to AM is substantially decreased following even a relatively brief exposure to AM of the same (or similar) rate.

(2) The function fitted to the signal modulation depth at threshold plotted against a linear delay between the masker and signal AM reveals exponential recovery from forward masking with the time constants falling in the range between 56 and 111 ms, for the four subjects used in the study.

(3) The AM forward masking effect shows tuning. The tuning appears to be broader than that observed in the simultaneous masking of AM. The broader tuning might result from the unavailability of cues based on local temporal features and modulation distortion products that may facilitate the detection of signal AM in simultaneous masking. The tuning does not depend on the signal modulation rate over a range of rates studied (20–80 Hz).

(4) The tuning is consistent with processing AM by rate-selective modulation channels and with a notion that neural activity in such channels adapts or persists after the offset of the AM that has stimulated the channel.

(5) For a single-rate AM, a 50% AM produces less forward masking of the same rate than a 100% AM, suggesting a monotonic relationship between the power of the masker and the amount of masking. Surprisingly, combining two equally effective masker AMs does not produce an increase in the amount of forward masking relative to that observed for either of the component maskers alone. This result is inconsistent with predictions by an envelope power spectrum model.

ACKNOWLEDGMENTS

We thank the editor Gerald Kidd, Jr., and two anonymous reviewers for their helpful comments on an earlier version of this paper. This work was supported by NIH/NIDCD Grant No. DC00683 and Grant No. DC006804.

[1] Masked-threshold patterns were first measured at a delay of 20-ms for all listeners and all rates of S_{AM}. When for the same (or similar) rates of S_{AM} and M_{AM}, the masked threshold could not be measured at the 20-ms delay due to the listener's inability to detect the signal at a 0-dB modulation depth, a delay longer by 10 ms was used. This was done because attenuation functions could not be derived without estimating the peak of the masking function. For those short delays, data points (not shown) were obtained for masker rates that did not produce the maximum amount of masking. A visual comparison of the shape of the attenuation functions measured at 20 ms with those measured at 30 or 40 ms shown in Fig. 4, suggests a parallel shift of the attenuation functions with increasing masker-signal delay. A parallel shift implies that the rate of recovery is independent of the stimulating modulation rate.

Bacon, S. P., and Grantham, W. D. (**1989**). "Modulation masking: Effects of modulation frequency, depth, and phase," J. Acoust. Soc. Am. **85**, 1575–2580.

Bartlett, E. L., and Wang, X. (**2005**). "Long-lasting modulation by stimulus context in primate auditory cortex," J. Neurophysiol. **94**, 83–104.

Cooper, N. P., Robertson, D., and Yates, G. K. (**1993**). "Cochlear nerve fiber responses to amplitude-modulated stimuli: variations with spontaneous rate and other response characteristics," J. Neurophysiol. **70**, 370–386.

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

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

Ewert, S. D., and Dau, T. (**2000**). "Characterizing frequency selectivity for envelope fluctuations," J. Acoust. Soc. Am. **108**, 1181–1196.

Ewert, S. D., Verhey, J. L., and Dau, T. (**2002**). "Spectro-temporal processing in the envelope-frequency domain," J. Acoust. Soc. Am. **112**, 2921–2931.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (**1990**). "Encoding of amplitude modulation in the gerbil cochlear nucleus. II. Possible neural mechanisms," Hear. Res. **44**, 123–142.

Füllgrabe, C., Moore, B. C. J., Ewert, S. D., Sheft, S., and Lorenzi, C. (**2005**). "Modulation masking produced by second-order modulators," J. Acoust. Soc. Am. **117**, 2158–2168.

Grimault, N., Bacon, S., and Micheyl, S. (**2002**). "Auditory stream segregation on the basis of amplitude-modulation rate," J. Acoust. Soc. Am. **111**, 1340–1348.

Hedrick, M. S., and Carney, A. E. (**1997**). "Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants," J. Speech Lang. Hear. Res. **40**, 1445–1457.

Hedrick, M. S., and Jesteadt, W. (**1996**). "Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing-impaired listeners," J. Acoust. Soc. Am. **100**, 3398–3407.

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

Joris, P. X., Schreiner, C. E., and Rees, A. (**2004**). "Neural processing of amplitude-modulated sounds," Physiol. Rev. **84**, 541–577.

Kay, R. H., and Matthews, D. R. (**1972**). "On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency modulated tones," J. Physiol. (London) **225**, 625–667.

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

Krishna, B. S., and Semple, M. N. (**2000**). "Auditory temporal processing: Responses to sinusoidally amplitude-modulated tones in the Inferior Colliculus," J. Neurophysiol. **84**, 255–273.

Kwon, B. J., and Turner, C. W. (**2001**). "Consonant identification under maskers with sinusoidal modulations: Masking release or modulation interference," J. Acoust. Soc. Am. **110**, 1130–1140.

Langner, G., and Schreiner, C. (**1988**). "Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms," J. Neurophysiol. **60**, 1799–1822.

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

Liang, L., Lu, T., and Wang, X. (**2002**). "Neural representations of sinusoidal amplitude and frequency modulations in the auditory cortex of awake primates," J. Neurophysiol. **87**, 2237–2261.

Lorenzi, C. (2004). personal communication.

Middlebrooks, J. C. (2005). "Transmission of temporal information from a cochlear implant to the auditory cortex," XXVIII Midwinter Research Meeting of ARO, Abstract No. 253.

Moody, D. B., Cole, D., Davidson, L. M., and Stebbins, W. C. (**1984**). "Evidence for a reappraisal of the psychophysical selective adaptation paradigm," J. Acoust. Soc. Am. **76**, 1076–1079.

Moore, B. C. J., and Glasberg, B. R. (**1982**). "Contralateral and ipsilateral cuing in forward masking," J. Acoust. Soc. Am. **71**, 942–945.

Moore, B. C. J., Sek, A., and Glasberg, B. R. (**1999**). "Modulation masking produced by beating modulators," J. Acoust. Soc. Am. **106**, 908–918.

Neff, D. L. (**1985**). "Stimulus parameters governing confusion effects in forward masking," J. Acoust. Soc. Am. **78**, 1966–1976.

Neff, D. L. (**1986**). "Confusion effects with sinusoidal and narrow-band noise forward maskers," J. Acoust. Soc. Am. **79**, 1519–1529.

Nelson, D. A., and Donaldson, G. S. (**2002**). "Psychophysical recovery from pulse-train forward masking in electric hearing," J. Acoust. Soc. Am. **112**, 2932–2947.

Nelson, D. A., and Freyman, R. L. (**1987**). "Temporal resolution in sensorineural hearing-impaired listeners," J. Acoust. Soc. Am. **81**, 709–720.

Plack, C. J., and Oxenham, A. J. (**1998**). "Basilar-membrane nonlinearity and the growth of forward masking," J. Acoust. Soc. Am. **103**, 1598–1608.

Rees, A., and Møller, A. R. (**1983**). "Responses of neurons in the inferior colliculus of the rat to AM and FM tones," Hear. Res. **10**, 301–310.

Regan, D., and Tansley, B. W. (**1979**). "Selective adaptation to frequency-modulated tones: Evidence for an information-processing channel selectively sensitive to frequency changes," J. Acoust. Soc. Am. **65**, 1249–1257.

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

Richards, V. M., Buss, E., and Tian, L. (**1997**). "Effects of modulator phase for comodulation masking release and modulation detection interference," J. Acoust. Soc. Am. **102**, 468–476.

Roberts, B., Glasberg, B. R., and Moore, B. C. J. (**2002**). "Primitive stream segregation of tone sequences without differences in F0 or passband," J. Acoust. Soc. Am. **112**, 2074–2085.

Schulze, H., and Langner, G. (**1999**). "Auditory cortical responses to amplitude modulations with spectra above frequency receptive fields: evidence for wide spectral integration," J. Comp. Physiol. [A] **185**, 493–508.

Sek, A., and Moore, B. C. J. (**2002**). "Mechanisms of modulation gap detection," J. Acoust. Soc. Am. **111**, 2783–2792.

Sheft, S. (2000). "Adaptation to amplitude modulation," XXIII Midwinter Research Meeting of ARO, Abstract No. 5058.

Sheft, S., and Yost, W. A. (**1990**). "Temporal integration in amplitude modulation detection," J. Acoust. Soc. Am. **88**, 796–805.

Shofner, W. P., Sheft, S., and Guzman, S. J. (**1996**). "Responses of ventral cochlear nucleus units in the chinchilla to amplitude modulation by low-frequency two-tone complexes," J. Acoust. Soc. Am. **99**, 3592–3605.

Strickland, E. A., and Viemeister, N. F., (**1996**). "Cues for discrimination of envelopes," J. Acoust. Soc. Am. **99**, 3638–3646.

Takahashi, G. A., and Bacon, S. P., (**1992**). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," J. Speech Hear. Res. **35**, 1410–1421.

Tansley, B. W., and Suffield, J. B., (**1983**). "Time course of adaptation and recovery of channels selectively sensitive to frequency and amplitude modulation," J. Acoust. Soc. Am. **74**, 765–775.

van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P., (**1987**). "Speech waveform envelope cues for consonant recognition," J. Acoust. Soc. Am. **82**, 1152–1161.

Viemeister, N. F., (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.

Wakefield, G. H., and Viemeister, N. F., (**1984**). "Selective adaptation to linear frequency-modulated sweeps: evidence for direction-specific FM channels?," J. Acoust. Soc. Am. **75**, 1588–1592.

Wang, X., and Sachs, M. B., (**1993**). "Neural encoding of single-formant stimuli in the cat. I. Responses of auditory nerve fibers," J. Neurophysiol. **70**, 1054–1075.

Wojtczak, M., and Viemeister, N. F., (**2003**). "Suprathreshold effects of adaptation produced by amplitude modulation," J. Acoust. Soc. Am. **114**, 991–997.

Wright, B. A., and Dai, H., (**1998**). "Detection of sinusoidal amplitude modulation at unexpected rates," J. Acoust. Soc. Am. **104**, 2991–2996.