JARO
Journal of the Association for Research in Otolaryngology

Forward Masking in the Amplitude-Modulation Domain for Tone Carriers: Psychophysical Results and Physiological Correlates

Magdalena Wojtczak[1,6], Paul C. Nelson[2,3], Neal F. Viemeister[1], and Laurel H. Carney[3,4,5]

[1]Department of Psychology, University of Minnesota, 75 East River Rd, Minneapolis, MN 55455, USA

[2]Center for Hearing and Balance, Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA

[3]Institute for Sensory Research, Department of Biomedical and Chemical Engineering, Syracuse University, New York, NY, USA

[4]Department of Biomedical Engineering, University of Rochester, Rochester, NY 14642, USA

[5]Department of Neurobiology & Anatomy, University of Rochester, Rochester, NY 14642, USA

[6]Auditory Perception and Cognition Lab, Elliott Hall, N218, 75 East River Rd, Minneapolis, MN 55455, USA

ABSTRACT

Wojtczak and Viemeister (J Acoust Soc Am 118:3198–3210, 2005) demonstrated forward masking in the amplitude-modulation (AM) domain. The present study examined whether this effect has correlates in physiological responses to AM at the level of the auditory midbrain. The human psychophysical experiment used 40-Hz, 100% AM (masker AM) that was imposed on a 5.5-kHz carrier during the first 150 ms of its duration. The masker AM was followed by a 50-ms burst of AM of the same rate (signal AM) imposed on the same (uninterrupted) carrier, either immediately after the masker or with a delay. In the physiological experiment, single-unit extracellular recordings in the awake rabbit inferior colliculus (IC) were obtained for stimuli designed to be similar to the uninterrupted-carrier conditions used in the psychophysics. The masker AM was longer (500 ms compared with 150 ms in the psychophysical experiment), and the carrier and modulation rate were chosen based on each neuron's audio- and envelope-frequency selectivity. Based on the average discharge rates of the responses or on the temporal correlation between neural responses to masked and unmasked stimuli, only a small subset of the population of IC cells exhibited suppression of signal AM following the masker. In contrast, *changes* in the discharge rates between the temporal segments of the carrier immediately preceding the signal AM and during the signal AM varied as a function of masker-signal delay with a trend that matched the psychophysical results. Unless the physiological observations were caused by species differences, they suggest that stages of processing higher than the IC must be considered to account for the AM-processing time constants measured perceptually in humans.

Keywords: recovery from modulation masking, modulation coding in the inferior colliculus

Correspondence to: Magdalena Wojtczak · Auditory Perception and Cognition Lab · Elliott Hall, N218, 75 East River Rd, Minneapolis, MN 55455, USA. Telephone: +1-612-6257056; fax: +1-612-6263359; email: wojtc001@umn.edu

INTRODUCTION

Temporal envelopes play an important role in the identification of complex stimuli such as speech and in the segregation of target sounds from backgrounds with overlapping spectra (Grimault et al. 2002). Severely degrading information contained in the temporal envelope with the fine structure information

preserved leads to a greater reduction in speech intelligibility than eliminating the fine structure of the speech sound by replacing it with random noise while preserving the envelope (e.g., Drullman 1995; Noordhoek and Drullman 1997; Gilbert et al. 2007). Envelopes of speech and other natural sounds are complex waveforms and can be represented by a sum of multiple sinusoidal modulations with different rates. How these modulation components are processed in the auditory system affects the temporal information available to the listener.

Psychophysical studies investigating modulation processing have shown that sinusoidal components in the modulation domain can interact and produce effects analogous to those resulting from interactions of audio frequencies within an auditory filter. One example is masking between modulation components imposed on the same carrier (Bacon and Grantham 1989; Houtgast 1989; Takahashi and Bacon 1992). Modulation masking, which exhibits relatively broad tuning, inspired models of envelope processing comprised of banks of modulation-frequency selective filters (Dau et al. 1997a, b; Ewert and Dau 2000; Ewert et al. 2002). The models typically are based on the assumption that a modulation filter bank exists for each peripheral auditory channel.

Another intriguing feature of perceptual AM processing is forward masking in the modulation domain (Wojtczak and Viemeister 2005). Using noise carriers, Wojtczak and Viemeister showed that a relatively short burst of modulation (150-ms masker AM) could make an ensuing burst of modulation (50-ms signal AM) imposed on the same carrier undetectable, for signal modulation depths that were easily detected when the preceding envelope was flat. This AM forward-masking paradigm may provide a more accurate psychophysical measure of tuning to AM than traditional simultaneous presentation. By presenting the masker and signal AM sequentially rather than simultaneously, any cues resulting from interactions between different AM components, such as local temporal features (Strickland and Viemeister 1996) or beats between the modulation components (Ewert et al. 2002), are not available. The authors speculated that the effect might reflect adaptation of AM rate-tuned neural responses, but the physiological basis for forward masking in the AM domain has not been experimentally explored.

Several known features of responses in the inferior colliculus (IC) make it a logical site to begin the search for neural correlates of the phenomenon. First, many IC neurons respond strongly only over a limited range of AM frequencies (e.g., Langner and Schreiner 1988; Krishna and Semple 2000; Joris et al. 2004; Nelson and Carney 2007) in a way that at least qualitatively mirrors psychophysical tuning estimates (Houtgast 1989; Ewert and Dau 2000; Wojtczak and Viemeister 2005). This rate tuning to AM is not present in the responses of auditory-nerve fibers (Joris and Yin 1992) but is instead centrally generated through a mixture of inherited processing from lower levels (e.g., Batra 2006) and further refinement within the IC itself (Tan and Borst 2007). Another emergent property of neural responses along the auditory pathway is their robust temporal context dependence. At the level of the IC, many cells exhibit strong suppression of responses to short tones by a preceding masker that is directly comparable to classical psychophysical forward masking (Nelson et al. 2009).

Given these apparent similarities between IC physiology and psychoacoustics, it was hypothesized that the parallels would also extend to the forward masking of AM. This study used tonal carriers to compare the time course of forward masking of AM observed psychophysically to that in neural responses recorded from the IC of the awake rabbit. If adaptation were involved, the prediction would be that responses to a brief probe AM should be reduced when preceded by an appropriately chosen AM masker and that this suppression should be strongest when the masker-probe delay is shortest. The data presented here do not support this simple hypothesis. However, an alternative response metric that compares responses immediately before and during the probe AM does vary systematically with delay. A downstream change-detecting observer with IC neurons as its inputs would predict psychophysical forward masking of AM. Overall, the IC recordings suggest that the psychophysical effect is probably mediated by relatively high-level neural processing, and they point out a limitation in the interpretation of IC neurons as physiological implementations of perceptual modulation filters.

PSYCHOPHYSICS: FORWARD MASKING OF AMPLITUDE MODULATION

Wojtczak and Viemeister (2005) measured recovery from forward masking in the amplitude modulation domain using noise carriers. Inherent envelope fluctuations in noise carriers could affect the amount of masking and the rate of recovery. In this study, the measurements of recovery from AM forward masking were extended to tonal carriers to evaluate the robustness of the effect to the type of carrier. Although one previous study showed forward masking in the AM domain for a high-frequency tonal carrier, the measurements were performed at only one short masker-signal delay, and thus, the rate of recovery cannot be estimated from those data (Moore et al. 2009). The use of tonal carriers was also motivated by the intended comparisons between the psychophysical data and physiological responses from the awake

rabbit IC. Tonal carriers were used to characterize basic properties of responses to AM in the rabbit IC in the study of Nelson and Carney (2007). The specific psychophysical task was similar to that used in the study of Wojtczak and Viemeister (2005).

Stimuli and procedure

Detection of the signal, a 50-ms burst of AM (S_{AM}), was measured as a function of the delay between the offset of a 150-ms masker AM (M_{AM}) and the onset of the S_{AM}. The M_{AM} and the S_{AM} had the same modulation rate of 40 Hz and were imposed on the same carrier, as schematically illustrated in Figure 1. The carrier was a 5.5-kHz tone with a duration of 500 ms, including 5-ms raised-cosine onset/offset ramps. The carrier was presented at 70 dB SPL. The M_{AM} had a modulation depth of 0 dB (defined in units of 20 log m, where m is the modulation index) and started with the onset of the carrier. Detection of the S_{AM} was measured at delays chosen at random for each run from a range

between 0 and 210 ms. The M_{AM} and the S_{AM} started at a 0-rad (sine) phase.

To obtain a reference for the amount of forward masking, detection of S_{AM} was measured without the preceding M_{AM} (i.e., with an unmodulated carrier preceding the signal, Fig. 1A). The unmasked threshold was measured for three different temporal positions of the S_{AM} within the carrier. The positions corresponded to the delays from carrier onset of 150, 240, and 360 ms (i.e., the delays of 0, 90, and 210 ms from the point where the M_{AM} would have ended if it had been present).

Detection thresholds were measured using an adaptive three-interval forced-choice procedure tracking the 79.4% correct point on the psychometric function (Levitt 1971). In the reference (control) condition (Fig. 1A), two of the three intervals contained an unmodulated carrier and the third interval, chosen at random on each trial, contained S_{AM}. In the test condition (Fig. 1B), two of the three intervals contained the same carrier but with the M_{AM} during the initial 150 ms, and the third interval, chosen at random on each trial, also contained S_{AM}.
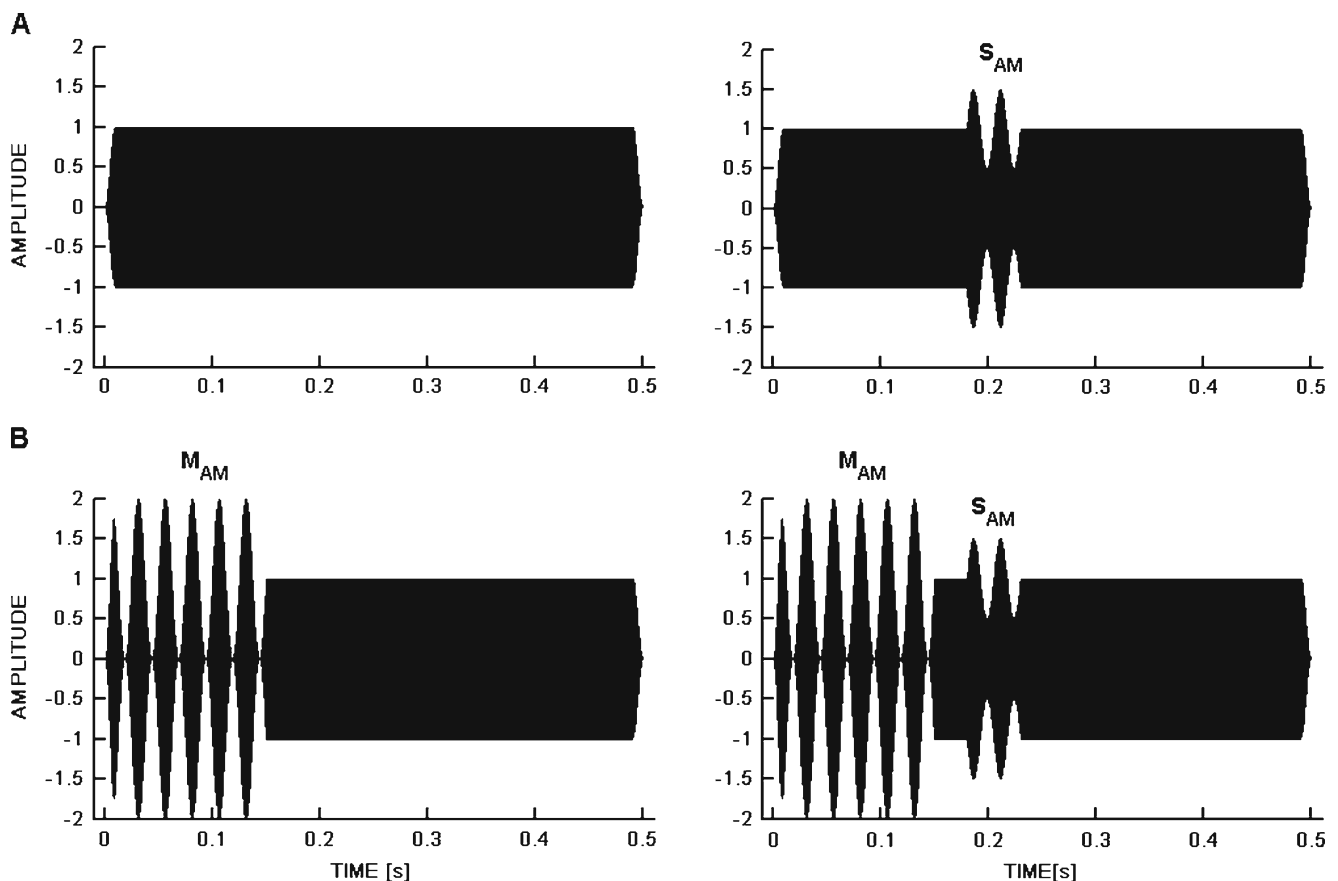

FIG. 1. Schematic illustration of the stimuli in the reference condition (**A**) and the test condition (**B**). The *left and right panels* show the stimulus in the non-signal and signal intervals, respectively. The delay between the masker (M_{AM}) and the signal (S_{AM}) in the example shown is 30 ms.

The intervals were marked by lights on a computer screen and were separated by 300-ms silent intervals. At the beginning of each run, the modulation depth of the S_{AM} was set to -2 dB. The modulation depth was decreased by 2 dB after three consecutive correct signal detections and increased by the same step after one incorrect response until four reversals were obtained. The step size was reduced to 1 dB for the subsequent eight reversals. Visual feedback was provided after each trial, following the listener's response. A run terminated after a total of 12 reversals was reached. Threshold signal modulation from a single run was calculated by averaging the modulation depths of S_{AM} at the last eight reversals. Four to six thresholds from single runs were averaged to obtain the final threshold estimate for each condition. When the adaptive procedure called for a modulation depth greater than 0 dB (modulation index $m>1$), the modulation depth was reset to 0 dB and the run continued. However, all of the runs in which this happened were excluded from further analysis. This approach was chosen over aborting the runs to allow for practice. For conditions in which the modulation depth called for exceeded 0 dB in three runs for a given subject, threshold was deemed immeasurable. In conditions for which this happened on fewer than three runs, a total of six runs were completed and the final estimate of threshold was calculated based on the runs for which the modulation depth did not exceed 0 dB during the adaptive tracking. Because these threshold estimates were biased, they were distinguished by using different symbol colors (as described below).

The stimuli were generated on a PC and played using a 24-bit sound card (Echo/Gina 24/96) with a sampling rate of 44.1 kHz and were routed to an earphone of a Sony MDR-V6 headset for monaural presentation to the left ear.

Listeners

Five listeners with normal hearing participated in the experiment. Their hearing thresholds were below 15 dB HL at audiometric frequencies from 0.25 to 8 kHz. The listeners were paid an hourly wage for their participation. All the listeners except one (S2, the first author) were naïve and were given at least 4 h of practice before data collection began. The listeners provided informed consent prior to their participation, and the protocol for the study was approved by the Institutional Review Board at the University of Minnesota.

Results

Data for individual listeners are shown in separate panels of Figure 2. The three thresholds for detecting the S_{AM} without the preceding M_{AM} are shown by filled triangles. A repeated-measures ANOVA showed that the unmasked threshold did not significantly depend on the delay of S_{AM} from the carrier onset ($F(1.9, 7.5)=2.0$, $p=0.201$); therefore, the three thresholds for each listener were averaged to provide the reference unmasked threshold (T_{unm}) represented by the dashed red line. The presence of the M_{AM} raised threshold for detecting the S_{AM} over the
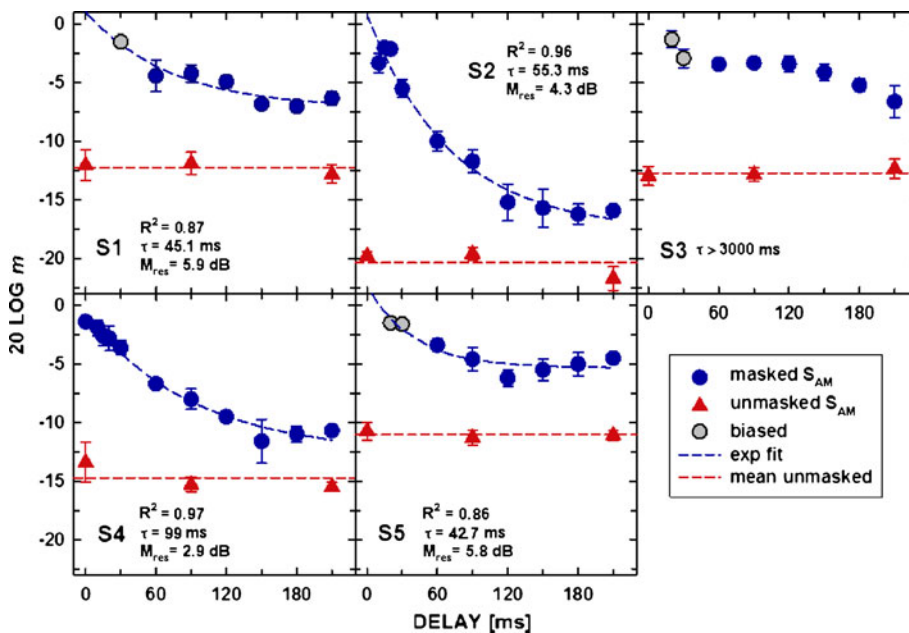

FIG. 2. Masked thresholds for detecting S_{AM} plotted as a function of $M_{AM}-S_{AM}$ delay (*blue symbols*). Thresholds measured without a masker are shown by *red triangles*. The *red dashed line* shows the mean unmasked threshold (T_{unm}), and the *blue dashed line* shows an exponential function fit to the data. The legend shows the parameters producing the best least-squares fit. The *gray symbols* represent thresholds that are biased by excluding the runs during which the listeners could not detect 100% S_{AM}.

entire range of delays used ($\leq$210 ms). For the shortest delay (0 ms), only one listener (S4) was able to detect the S_{AM} without the adaptive tracking calling for a modulation depth greater than the maximum allowable depth of 0 dB (100% modulation). For all the other listeners, the adaptive tracking called for a modulation depth greater than 0 dB in at least three runs, and the threshold was deemed immeasurable. Three listeners (S1, S3, and S5) could not detect the S_{AM} with a 100% (0-dB) modulation for other brief delays, as evidenced by missing symbols corresponding to the delays up to 20–30 ms. Data shown by the gray circles represent thresholds for delays for which the threshold for detecting the S_{AM} could be measured in the majority of runs but not in all of them. These threshold estimates are therefore biased by excluding the runs in which the modulation depth during adaptive tracking exceeded 0 dB. For longer delays, threshold for detecting the S_{AM} decreased progressively with increasing delay up to about 120–150 ms and then reached an asymptotic residual masking of about 5 dB above T_{unm}. The exception was S3 whose data showed very slow recovery with no clear asymptote for the range of delays used. The patterns of the results were qualitatively similar to those found by Wojtczak and Viemeister (2005) for a noise carrier. For two listeners (S1 and S2), thresholds were measured for two additional delays of 360 and 510 ms (data not shown). For S2, threshold recovered to the unmasked value at the 510-ms delay, whereas for S1, there was about 2 dB of residual masking at that delay. Since the threshold changed very slowly over the delays longer than 210 ms and using those delays required substantially increasing the duration of each carrier and consequently each observation interval, the measurements presented in this study were limited to delays $\leq$210 ms.

Recovery functions, converted to masking curves by subtracting T_{unm} from each threshold, were fitted with an exponential function, defined by

$$M(d) = ae^{\frac{-d}{\tau}} + M_{res},$$

where $M(d)$ denotes masked threshold as a function of masker-signal delay (d), a is a parameter, τ is the estimated recovery time constant, and M_{res} is an asymptotic (residual) amount of masking. The exponential function was fitted by minimizing the sum of squared deviations between the data and the fit using 200 iterations. The time constants, τ; the residual masking, M_{res}; and the proportion of variance accounted for by the fitted functions (R^2) are shown in the legend in each panel. The exponential function did not converge on a solution, and these parameters could not be determined for listener S3 within 200 iterations. Increasing the number of iterations led to a fit that produced an estimated recovery time constant $\tau>3$ s. Because of the anomalous result, the data for S3 were excluded from further analysis. The average recovery time constant for the remaining four listeners was 60.5 ms. This time constant was shorter than the average time constant of 97 ms calculated using the results for the noise carrier from the study of Wojtczak and Viemeister (2005), but an independent sample t test showed that the recovery times for the two types of carriers were not significantly different ($t(6)=1.93$, $p=0.101$). The data and the fitted functions reveal substantial residual masking at long masker-signal delays (see M_{res} values in each panel of Fig. 2). Significantly less residual masking was observed for the noise carrier used by Wojtczak and Viemeister. The average residual masking for the tonal carrier was 4.8 dB, whereas the residual masking for the noise carrier was 1.8 dB. An independent sample t test showed that the difference in residual masking for the two types of carrier was statistically significant ($t(6)=3.66$, $p=0.011$).

Discussion of psychophysical findings

Forward masking in the AM domain was observed for tonal carriers. The data in Figure 2 show that a relatively short burst of the maximum (100%) AM cannot be consistently detected for bursts of modulation with offset-to-offset intervals as long as 70–80 ms following a longer burst of the same rate AM. These time intervals are relevant to the perception of natural sounds, such as speech. The relevant information in those sounds is carried by suprathreshold changes in amplitude or frequency over time. The data presented here show that suprathreshold temporal modulation can be completely masked by preceding modulation, and thus, the effect may limit the amount of information available in the temporal envelope of ongoing dynamically varying natural stimuli.

The time course of recovery from AM forward masking observed for the tonal carrier was on average similar to that for the noise carrier in the study by Wojtczak and Viemeister (2005). The intriguing result is the relatively large residual AM masking that may extend to much longer delays than those used in our experiment. The residual masking may reflect cumulative adaptation due to the repeated presentation of the masker modulation over trials within each run (Gutschalk et al. 2008).

Wojtczak and Viemeister (2005) and Moore et al. (2009) showed that the forward masking of AM exhibits tuning, suggesting that the physiological mechanism underlying the forward masking operates at or above the stage of envelope processing where rate-selective tuning in neural responses to AM is first observed. The following experiment and analyses

represent the first attempt to find physiological correlates of the forward masking in the modulation domain observed in psychophysics.

PHYSIOLOGY

Methods

A detailed description of animal preparation and recording methods, along with a characterization of basic AM response properties in the awake rabbit IC, can be found in Nelson and Carney (2007). A brief account is provided here, with an emphasis on differences in stimuli and response analyses used in the AM forward-masking paradigm.

Animal preparation and recording. Access to the midbrain of two female Dutch-belted rabbits (*Oryctolagus cuniculus*) was provided by a small craniotomy that was positioned within a stainless steel recording chamber used as a base for the electrode micropositioner. Animals were restrained using a headbar and a blanket that was wrapped snugly around the body. Recording sessions typically lasted for 2 h and were terminated if the animal showed signs of discomfort. All procedures were approved by the Syracuse University Institutional Animal Care and Use Committee and conformed to the National Institutes of Health standards.

Well-isolated single-unit extracellular responses were acquired using glass-coated tungsten electrodes, and Schmidt-triggered action potentials were timed with respect to the onset of the acoustic stimulus. Entry into the physiologically defined central nucleus of the IC was marked by the presence of strongly sound-driven background activity and an orderly increasing progression of best frequency (BF).

Acoustic stimuli and response analysis. Sounds were presented via a closed and calibrated acoustic system, either monaurally to the ear contralateral to the recording site ($n=25$) or diotically to both ears ($n=11$); the configuration that elicited higher driven rates to 500-ms broadband noise bursts was used for the remainder of the protocol. All test stimuli used BF tone carriers presented at an SPL typically 20–40 dB above the neuron's threshold. Figure 3A, B shows an example of one neuron's pure-tone response profiles and the resulting BF and SPL used with the AM stimuli.

To determine the cell's AM tuning characteristics, a modulation transfer function (MTF) was obtained (e.g., Fig. 3C) by presenting two or three repetitions of 2-s duration, fully modulated S_{AM} tones at 15 modulation frequencies logarithmically spaced between 2 and 311 Hz. The AM frequency that resulted in the highest value of synchronized rate, defined as the product of average rate and vector strength, referred to as the best modulation frequency
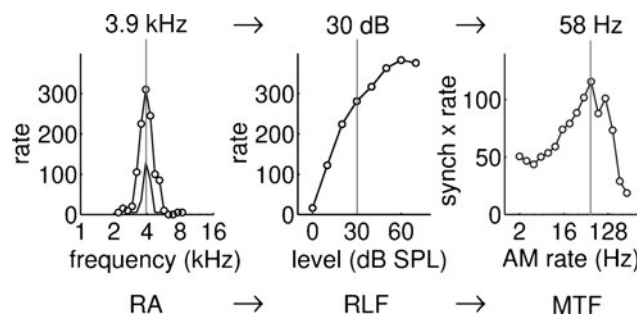

FIG. 3. Progression of stimulus parameter selection leading up to physiological AM forward-masking paradigm. The carrier frequency was set to the neuron's characteristic frequency based on a response area measured 10 and 40 dB above the audio-visually estimated threshold (*RA*; *left*). Overall level was set to a point 20–40 dB above threshold, usually on the initial ascending portion of the rate-level function (*RLF*; *middle*). Modulation frequency (for both masker and probe) was chosen to match the peak in the synchronized rate modulation transfer function (*MTF*; *right*).

(BMF) was chosen as the modulation rate for the remainder of the protocol. In two neurons with complex MTFs, a secondary peak in the synchronized rate function was chosen to accommodate more cycles of the signal AM within a short signal presentation time.

The response dependence on signal modulation depth was mapped out by measuring a modulation depth function (MDF) at eight modulation depths, from −35 dB (1.8%) to 0 dB (100%) in 5-dB steps, for a modulation rate corresponding to the BMF.

The AM forward masking stimuli used in the physiology were modeled after the psychophysical stimuli, with parameters designed to maximize the potential impact of the masker modulation. The masker duration was extended to 500 ms. The masker duration was longer than the 150-ms duration in the psychophysics because a set of pilot results (unpublished) from three human listeners obtained with a noise carrier showed a trend for an increased amount of masking for 500- compared with 150-ms M_{AM} at masker-signal delays up to 60–90 ms. The modulation rates of masker and signal were identical and set equal to the BMF. A short signal modulation, S_{AM}, was embedded within the 500-ms steady carrier that immediately followed the masker. The S_{AM} duration was at least 50 ms; the exact duration varied somewhat to accommodate an integer number of modulation cycles. This requirement led to a range of effective signal durations from 50 to 70 ms depending on the chosen stimulus AM rates. Both $M_{AM} - S_{AM}$ delay and S_{AM} modulation depth were systematically varied; all neurons were tested with at least one delay for a signal modulation depth of 0 dB (100% modulation). The delays between the M_{AM} offset and S_{AM} onset were 12.5, 25, 50, 100, and 200 ms, although some cells were

not held long enough to record responses for all five delays. A control condition, in which the carrier was unmodulated both before and after the S_{AM} (i.e., no masker AM was used), was also included. In this condition, the signal was temporally centered in the carrier. Each 1-s sound was repeated 20 times with a minimum inter-stimulus interval of 1 s. In the control condition, the carrier onsets and offsets were gated with 50-ms cosine-squared ramps. Onsets with the M_{AM} were determined by the modulation period (started in sine phase) whereas offsets had the 50-ms ramps. The S_{AM} was physically identical in the control and masked conditions, and no additional windowing was applied to this modulation.

To quantify the effects of the presence of the masker on the response to the S_{AM}, a 70-ms analysis window was defined surrounding the response to the S_{AM}, starting 8 ms after its onset to compensate for absolute response latencies. This delay value was based on the minimum first-spike latencies observed in a larger population of rabbit IC neurons (Nelson and Carney 2007). No attempt was made to compensate differently for individual neurons, since ongoing latencies can be different from onset latencies in the IC (e.g., Chase and Young 2006), making any definition somewhat arbitrary. The chosen window was consistent with the temporal location of the responses based on visual inspection of peristimulus time (PST) histograms.

Average spike counts were compared in the control and masked conditions, and the difference in average firing rates and across-repetition variability of the rate observations were jointly taken into account by computing a d-prime metric (Sakitt 1973) from signal detection theory:

$$d' = \frac{|\text{rate}_{\text{masked}}| - |\text{rate}_{\text{control}}|}{\sqrt{\sigma(\text{rate}_{\text{masked}}) \, * \, \sigma(\text{rate}_{\text{control}})}}$$

For rate distributions with equal variances, an equivalent formulation replaces the denominator with the arithmetic mean of the standard deviations of the two distributions (instead of the geometric mean used here). Note that this quantity can be positive or negative; positive values indicate higher spike counts in the masked condition than in the control case, and negative values suggest a suppression of the masked response compared to the control response. A negative value of d' would be broadly consistent with adaptation (i.e., suppression of the response to the S_{AM}) as the mechanism underlying the forward masking observed in the psychophysics. A similar metric was used to quantify differences between other aspects of the neural responses, such as variations in rate during the S_{AM} in the control and test conditions and rate changes over the course of the M_{AM} presentation.

Results

Baseline response characteristics. Neural responses to the test stimuli were recorded in 36 units from within the physiologically defined central nucleus of the awake rabbit IC. Modulation waveforms were imposed on BF tone carriers, ranging in frequency from 330 Hz to 17.3 kHz (median BF=3.54 kHz), presented at an overall SPL 20–40 dB above the neuron's threshold (Fig. 3). The modulation rate of the M_{AM} and S_{AM} was set equal to each cell's BMF; across the population, the BMF varied from 16 to 310 Hz (the distribution is shown in Fig. 5).

Note that a peak in the synchronized rate MTF did not always correspond to a peak in the average rate MTF; in five neurons, the firing rate dropped when the tone was modulated, and in four cells, the average rate was independent of the signal's modulation depth. The remaining 27 neurons had monotonically increasing rate MDFs, although most functions did not begin to increase until relatively high modulation depths were used: The average rate-based neural detection threshold among the 32 units with significant changes in rate was −6.8 ± 6.0 dB (20 log m±1 SD). This group of cells is representative of a larger sample of neurons with similar AM sensitivity characteristics (Nelson and Carney 2007).

Because many neurons did not increase their firing rate until the stimulus was strongly modulated even in unmasked (control) conditions, the S_{AM} modulation depth was initially set to 0 dB (m=1), and the delay between M_{AM} offset and S_{AM} onset was varied over a range from 12.5 to 200 ms. A subset of neurons was also tested with lower S_{AM} modulation depths (usually −5 or −10 dB).

Diverse but subtle masker effects on probe rate responses. To quantify the effect of modulating the tonal carrier in the period leading up to the S_{AM} presentation, firing rates were compared during the S_{AM} response in control (without M_{AM}) and masked (with M_{AM}) conditions. The physical S_{AM}, along with the 12.5-ms unmodulated period preceding it, was identical in the two cases. Responses of three example cells are shown in Figure 4A–F with (blue) and without (red) M_{AM}. The majority of neurons in the sample (21 out of 36) had similar responses to neuron 1 in that the response to the S_{AM} was *not* strongly affected by the masker modulation ($|d'|\leq1$). This was true despite the fact that the M_{AM} was a much more effective driver of the neuron than an equal-SPL unmodulated tone (compare the blue and red functions over the time interval (−500, 0)). A short time window PST histogram surrounding the time of S_{AM} presentation (Fig. 4B) further reveals little change in the temporal firing pattern, either in terms of response modulation magnitude (i.e., synchrony) or phase.

A subset of neurons (eight out of 36) did exhibit varying degrees of S_{AM}-response suppression after the M_{AM}. For example, neuron 2 in Figure 4C, D clearly responded more strongly to the S_{AM} in the control condition than in the masked case, leading to a d' of -2.45. This suppressive effect was the strongest during the first several cycles of the 80-Hz S_{AM} and was substantially weaker for subsequent stimulation periods. Similar to neuron 1, the effect did not seem to substantially alter the timing of the phase-locked response. In contrast to the first two examples, neuron 3 (Fig. 4E, F) is representative of the seven cells that responded to the S_{AM} with more spikes when the masker modulation was present than when the carrier was unmodulated prior to the onset of the S_{AM}. One feature of this neuron's response is the buildup over time in response to the M_{AM} (Fig. 4E). The M_{AM} appeared to "prime" the cell's ability to respond strongly to the S_{AM}, while

an unmodulated carrier preceding the signal elicited few spikes and resulted in a relatively weak ("unprimed") S_{AM} response. Spike timing was also qualitatively different in the control and masked conditions; the timing may be another reflection of the temporal buildup pattern observed in this cell in response to sustained AM.

Average firing rates (± 1 SD) are shown in Figure 4G for control and masked S_{AM} responses across the sample of IC neurons. The cells are sorted from the lowest to highest value of d', and the gray shaded area highlights the combinations of masked (blue) and control (red) rate responses that led to d' values between -1 and $+1$. When a criterion d' value of ± 2 was used, six cells were still classified as "suppressed" ($d' < -2$), while only two cells maintained their "enhanced" label ($d' > +2$). Therefore, suppression was typically more robust than enhancement. Still, the overall tendency was for IC neurons to
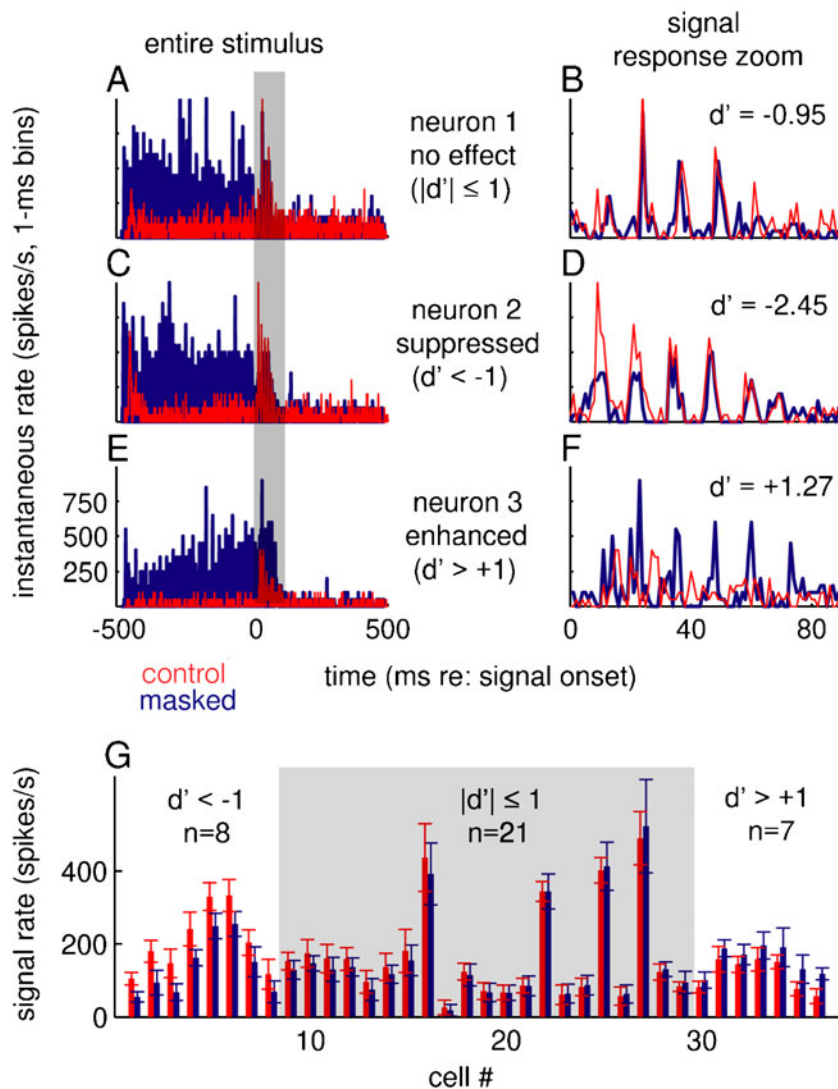

FIG. 4. Effects of M_{AM} on S_{AM} responses of IC neurons. The initial 500 ms of a BF tone carrier was either modulated by M_{AM} (*blue*, "masked") or unmodulated (*red*, "control"), and a 12.5-ms steady-state tone was present between the M_{AM} and the S_{AM} (the M_{AM} histograms were shifted to align the S_{AM} presentation times). A–F Example responses from three IC neurons. All three example neurons responded more strongly when the M_{AM} was present and had a best modulation frequency of 80 Hz, but the effect on the subsequent S_{AM} could be negligible (A, B), suppressive (C, D), or enhancing (E, F). Responses are shown for the entire 1-s stimulus (A, C, E), or for an 80-ms window beginning at the physical S_{AM} onset (B, D, F). d' values for the S_{AM}-response differences are included above the magnified PSTHs. G Average S_{AM} rates ($\pm$ SD) in the two conditions for the sample of neurons, sorted from the most negative to the most positive d' values.

exhibit rather small absolute differences between masked and control S_{AM} responses, even for a short (12.5-ms) delay between the masker and probe.

In parallel with the psychophysical measurements, the dependence of the physiological responses on the delay between the M_{AM} and S_{AM} was also examined. Several log-spaced delays (from 12.5 to 200 ms) were tested. As another way to visualize the overall IC response, across-neuron-averaged PST histograms (10-ms bins) are plotted in the top panels of Figure 5 for the 28 neurons for which 20 repetitions of the 12.5-, 50-, and 200-ms delay conditions were all recorded. The timing of the S_{AM} is aligned in each case; because the M_{AM} had a fixed duration (500 ms), responses to stimuli with longer delays extend to earlier points in time with respect to the S_{AM} onset. As expected from the examples discussed above, average S_{AM} responses in the 12.5-ms delay condition were modestly suppressed. A somewhat unexpected result was that this suppression persisted in the averaged PST histogram for the 50- and 200-ms delay conditions as well. This result was surprising because in the 200-ms delay case,
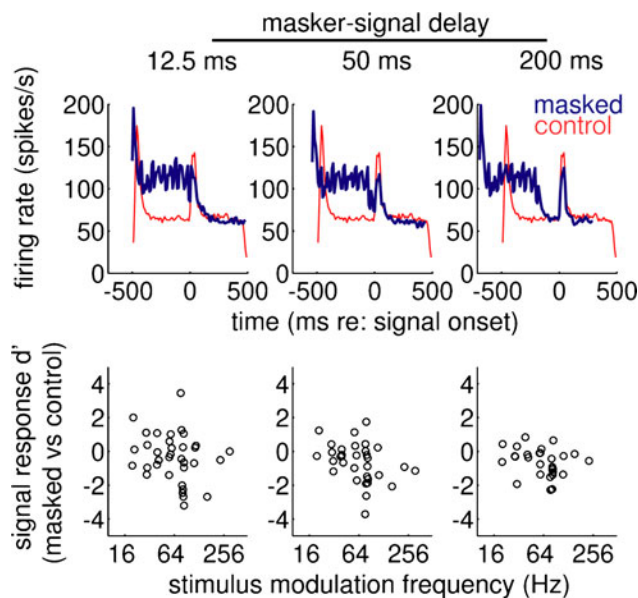
the response appears to have recovered to its baseline (control) level prior to the presentation of the S_{AM}.

To understand the origin of this apparent discrepancy, it is helpful to return to a cell-by-cell analysis of the S_{AM} responses. The lower panels of Figure 5 show S_{AM}-response d' (as defined above) plotted as a function of the stimulus modulation frequency for the same three delay values. As the delay is increased, the range of observed d' values is compressed. Strong positive values of d' (reflecting masker-induced enhancement of the probe response) are not observed in the 200-ms condition (Table 1). This result suggests that the residual suppression observed for longer delays in the upper panels of Figure 5 is influenced more by the absence of enhanced neurons than by the presence of neurons with strong, long-lasting suppressive effects. A small subsample of neurons ($n=8$) was also tested with an S_{AM} modulation depth of -5 dB ($m=0.56$) and/or -10 dB ($m=0.32$). In the -10-dB case, only 1/8 cells exhibited suppression ($d'<-1$); in the -5-dB condition, 1/4 neurons yielded a $d'<-1$.

Can masker responses predict S_{AM} suppression?. One possible mechanism for physiological forward suppression in the IC is the depletion of excitatory neurotransmitter. Two straightforward predictions can be tested to determine if simple firing-rate adaptation is consistent with the (rather small) suppressive effects that were observed. First, neurons that have a stronger preference for AM forward maskers (versus unmodulated carriers) should give rise to stronger suppressive effects. To quantify a neuron's preference for AM, firing rate distributions during the 500-ms M_{AM} and during the 500-ms unmodulated carrier (control) conditions were compared, and a d' statistic analogous to that used above was applied to describe S_{AM}-response differences. S_{AM}-response d' is plotted against this AM-preference metric in Figure 6A. As mentioned above, due to our stimulus parameter selection based on synchronized rate instead of average rate, not all neurons responded with more spikes to AM than to a pure tone, as shown by the symbols to the left of the vertical line. Even among those with a preference for AM, both suppression and enhancement were observed, and the change in the rate response depending on whether or not M_{AM} was present was not correlated with the S_{AM}-response d'. However, if only the lower-right quadrant of Figure 6A was considered (restricting the analysis to the neurons that showed both suppression and a preference for AM), a significant negative correlation ($r=-0.74$, $p<0.01$) was observed.

The second testable and intuitive prediction if an adaptation-like phenomenon was responsible for the effect is that the firing rates of cells that exhibit AM forward suppression should decrease over the course of the masker AM. This prediction was examined by computing another d' metric based on a comparison

FIG. 5. Summary of neural forward masking of AM across the sample of IC cells. *Top panels*: across-neuron average PST histograms (10-ms bins) for control (*red*) and AM-masked (*blue*) conditions for three values of the delay between the M_{AM} offset and the S_{AM} onset. Masked responses are aligned to the control condition S_{AM} onset time. Only neurons for which all three delay conditions were measured are included in the average histograms ($N=28$). *Bottom panels*: individual neuron S_{AM}-response d' statistics for the same three delay conditions as a function of the stimulus AM rate. A small amount of jitter was added to the x-axis values to facilitate visualization, since a discrete set of stimulus AM frequencies was used (based on each neuron's preferred modulation frequency). These two quantities (d' and modulation rate) were not strongly correlated with one another (only the 50-ms condition yielded a significant ($p=0.04$) negative ($r=-0.36$) correlation).

TABLE 1						
S_{AM}-response d' distribution statistics for short (12.5 ms) and long (200 ms) $M_{AM}-S_{AM}$ delays						
Delay	Number	Mean	Median	SD	Max	Min
12.5	36	−0.33	−0.25	1.39	+3.46	−3.20
200	28	−0.65	−0.49	0.82	+0.84	−2.28

between the first and last 100 ms of the M_{AM} response. Similar results were obtained for other window durations (data not shown). If adaptation were correlated with S_{AM} suppression, then large positive values of this measure should be associated with stronger suppression. Figure 6B shows that this was not the case: The two quantities were not significantly correlated with one another, even when the analysis was confined to cells that did adapt and showed suppression ($p>0.05$). To summarize, a simple neurotransmitter-depletion mechanism may underlie the forward-masked AM responses of some subset of IC neurons, but given the scatter of the data in the two panels of Figure 6B, it is clear that other mechanisms are in play.

Regardless of the underlying mechanism, the fact remains that it is difficult to reconcile the IC physiological responses with the psychophysics when the control and masked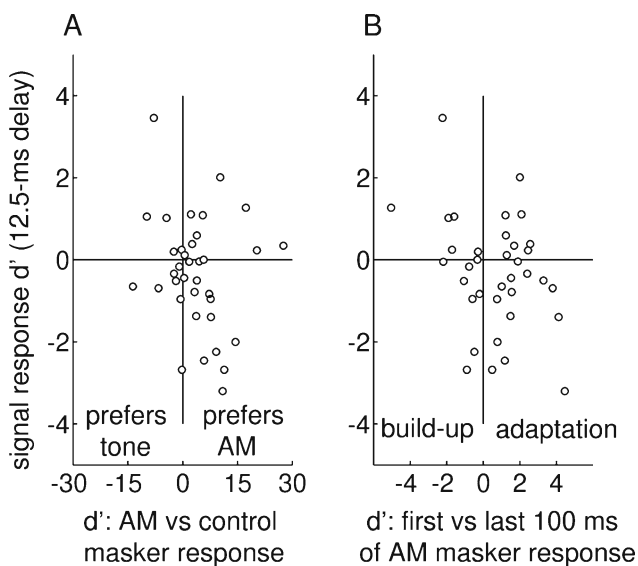 S_{AM} responses are compared with one another. Effects of modulating the portion of the carrier preceding S_{AM} were typically weak and did not systematically change with delay in a way that allows for straightforward predictions of the psychophysical data shown above. In the next section, a different analysis strategy is considered that compares the magnitude of local rate changes at the transition between the pure-tone delay period and the S_{AM} onset.

IC contrast effects predict delay-dependent AM forward masking. A qualitative inspection of the average PST histograms in Figure 5 leads to the observation that the "contrast" between the periods immediately prior to and during the S_{AM} response appears much more salient in the control and long-delay conditions than for shorter delays. To determine whether this was true in individual neurons (as opposed to the sample average PST histograms), single-cell firing rates were measured in two windows: during the 70-ms S_{AM}-response epoch and during the 70 ms immediately preceding the S_{AM} response (see "Methods"). Although, for simplicity, this analysis used a fixed window size, it does not imply a necessity for the assumption that a read-out mechanism at some higher level of the system must have a priori knowledge of the S_{AM} duration. The contrast could be quantified from the output of a sliding temporal integrator, analogous to the one used to account for forward masking in the domain of audio frequencies (e.g., Moore et al. 1988). As with the other response discriminability metrics, the rate differences in the two windows were normalized by their across-repetition variability, resulting in the new measure of contrast d'. Positive (negative) values of contrast d' arise when the S_{AM} response is larger (smaller) than the response for the pre-signal transition interval. Note that this metric does not require explicit comparisons across masked and unmasked responses; instead, a potential strategy for an efficient observer might be to choose the interval that yielded the largest contrast d' in a forced-choice task. This strategy is more analogous to the psychophysical task design, in which control and masked thresholds were measured in separate blocks of trials.

The resulting contrast d' values are shown in the panels of Figure 7, for the five tested masker-signal delays, along with the control condition. Portions of the average PST histograms from Figure 5 are shown above the corresponding stimulus conditions; the gray arrows schematize the magnitude of the average

FIG. 6. Differences in rates of masker (unmodulated versus M_{AM}) responses are poor predictors of variation across neurons in S_{AM}-response suppression. *Left*: S_{AM} d' using a 12.5-ms delay (as in the left panels of Fig. 5) as a function of the discriminability between overall firing rates elicited by the M_{AM} and unmodulated precursors. Positive (negative) values are obtained when the response to the M_{AM} is larger (smaller) than that to the unmodulated precursor. *Right*: ordinate values are identical to the *left panel* (i.e., probe response d'); abscissa values represent the differences in firing rates during the first and last 100 ms of the M_{AM} response. Positive values reflect a drop in rate over time; negative numbers result when the response builds up over the course of the M_{AM}. Neither metric (*left* or *right panel*) was significantly correlated with the S_{AM}-response d' ($p>0.05$).

contrast effect. Individual neural contrast d' values were tightly clustered around zero in the 12.5-ms delay condition. These distributions systematically spread toward more positive values as the delay was increased. With a 200-ms delay, the contrast d' distribution was qualitatively similar to that for the control. Strongly negative contrast d' values were rare (for six out of 160 observations, $d' < -2$) and did not typically vary monotonically with delay. In control conditions, two third of the sample (24 out of 36) responded with a contrast d' greater than unity.

Figure 8 summarizes the results illustrating the potential for IC responses to explain psychophysical AM forward masking based either on S_{AM}-response suppression or on changes in local contrast. The basic trend of interest is the recovery of perceptual sensitivity to the S_{AM} toward that for the control as the duration of the masker-signal gap is lengthened. In terms of proportions of cells exhibiting suppression, fewer cells would be expected to be suppressed when longer delays were used; by definition, no cells would exhibit suppression in the control condition.
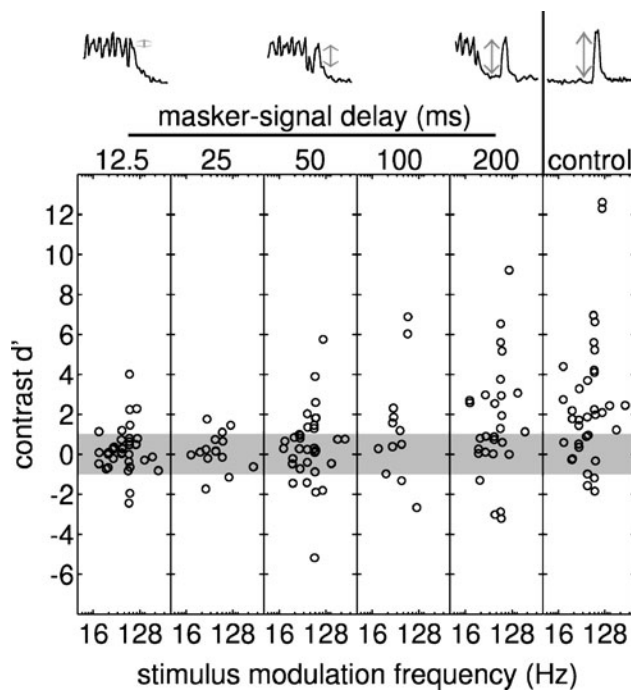
The expectation for the contrast function would be reversed, in a sense: More cells should exhibit significant contrast responses as the delay was increased. These expectations are schematized in the left panels of Figure 8 (the absolute proportions shown are somewhat arbitrary; the trends are not).

The data clearly support the contrast hypothesis more convincingly than the suppression hypothesis. The right column of Figure 8 shows the proportions of IC neurons in the sample with significant suppression (top) or contrast (bottom), for two absolute d' levels, 1 and 2. Using the more liberal d' criterion ($d' < -1$), the number of suppressed neurons actually *increased* for longer masker-signal delays. This prediction is the opposite of the expectation from the psychophysical data. When a more conservative d' criterion was used, the trend reversed, but less than 20% of the sample exhibited suppression, even with the shortest tested delay (12.5 ms). Overall, the presence of suppressed S_{AM} responses did not provide a robust prediction of the psychophysical effect. On the other hand, the proportion of neurons that yielded significant contrast responses increased systematically from a low level at short delays to near control levels with a delay of 200 ms, regardless of the criterion used to define d' significance. This distinction provides corroboration of the idea that perceptual AM forward masking is more likely to reflect local contrast at the level of IC neurons than absolute suppression of S_{AM} responses by the preceding masker AM.
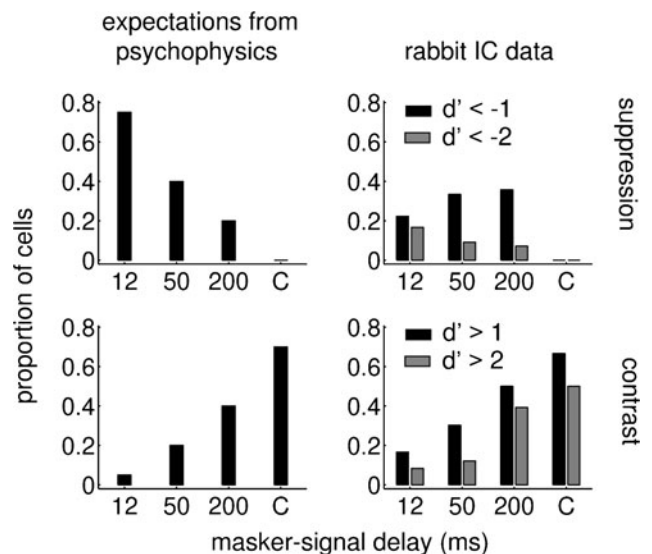

FIG. 7. As the masker-signal delay is increased, the response contrast between the periods immediately preceding the S_{AM} and during the S_{AM} increases systemically. Contrast d', defined as the discriminability of the response epoch 70 ms before the S_{AM} and the epoch during the S_{AM} response, is shown as a function of the stimulus AM frequency for five delay conditions and the control (unmodulated M_{AM}) condition. *Gray shading marks* contrast d' with absolute values less than unity. Average PST histograms (redrawn from Fig. Phys 3) are shown above their corresponding delay condition; a 650-ms period of the response is shown, from 150 to 800 ms post-M_{AM} onset. *Arrows* schematize the size of the local contrast upon presentation of the S_{AM}.

FIG. 8. Response contrast changes in a way broadly consistent with the psychophysics, while S_{AM}-response suppression does not. *Left panels*: expected trends for proportions of neurons with significantly modulated response suppression (*top*) and contrast (*bottom*) to qualitatively account for the psychophysical finding of delay-dependent forward masking of AM. *Right panels*: actual neural proportions for the same conditions. Two criterion d' values (1 and 2) were used to classify responses as significant (*black and gray bars*, respectively).

DISCUSSION

It is tempting to make comparisons between forward masking in the modulation domain and the more extensively studied forward masking in the audio-frequency domain. In both cases, exponential decay provides a reasonable characterization of the recovery process whereby masked threshold decays as the delay between the masker and signal increases. At delays longer than about 120–150 ms, the recovery from modulation-domain forward masking was very slow. The recovery time constants estimated from exponential functions fitted to the data were similar for the tonal carrier used in this study and the noise carrier used in the study of Wojtczak and Viemeister (2005). Interestingly, these time constants were in the range of those estimated from audio-frequency forward-masking data for listeners with little or no cochlear compression and thus, likely, reflected a pure higher-level recovery process (e.g., Nelson and Freyman 1987, for hearing-impaired listeners; Nelson and Donaldson 2002, for cochlear-implant users). This apparent similarity prompts a question of how far the analogy between the two forward-masking effects extends and whether they could have a similar underlying mechanism(s).

The mechanisms underlying IC physiological forward masking in the audio- and modulation-frequency domains appear to be fundamentally different. In the audio-frequency domain, responses to brief BF probe tones are almost universally suppressed by the presence of a BF forward masker (at least using diotic or contralateral stimulus presentation; e.g., Finlayson 1999), and this suppression often leads to neural detection threshold shifts that are similar to those measured in human psychophysics (Nelson et al. 2009). The results shown here demonstrate that when the experiment is translated into the AM-frequency domain, IC neurons do not tend to show strong suppression of signal responses. These parallel findings can be interpreted in the framework of a popular psychophysical model used to account for non-simultaneous masking effects in the audio-frequency domain. The temporal window model (e.g., Moore et al. 1988) assumes a smoothing stage followed by a decision device that searches for transient fluctuations in its output. Forward masking is assumed to be caused by some form of integration rather than adaptation or suppression. The physiological results support the existence of such a mechanism at the level of the IC in the modulation-frequency domain, but not in the audio-frequency domain, where response suppression is observed and is sufficient to predict perceptual forward masking (Nelson et al. 2009). The contrast metric proposed here to explain AM forward masking is essentially a simplified version of the temporal window model, with a coarse low-pass filtering stage (when firing rates are averaged in 70-ms bins before and during the S_{AM}), preceding a decision statistic that is based on changes in the smoothed local firing rates.

Most of the data presentation here was focused on a description of the average firing rates of IC neurons during (and before) S_{AM} stimulation. If more precise timing information is available and used by the system to detect modulation, it is possible that the forward M_{AM} could leave overall rates unaffected while disrupting the envelope synchronization (or some other temporal pattern) within the probe response. There were no systematic trends apparent in the histograms that supported this idea. A correlation analysis that compared control and masked responses (data not shown) did not show orderly changes in correlation coefficients as the delay between the masker and probe was varied, suggesting that fine temporal patterns were not affected by the masker in a way consistent with the psychophysical recovery from forward masking of AM.

The general role of IC response synchronization in the encoding of AM remains an open question. Unlike some neurons in auditory cortex (e.g., Bendor and Wang 2007), essentially all cells in the central nucleus of the IC are capable of following envelope fluctuations to some extent (Joris et al. 2004). Typically, such synchrony emerges at modulation depths lower than those required to elicit an increase in overall rate (Nelson and Carney 2007). For example, among the neurons in this study, the smallest modulation depth that resulted in significant synchrony was on average −17 dB, while the average neural detection threshold based on firing rate was −7 dB. The question of whether the animal uses the information available in temporal firing patterns at lower modulation depths is complicated somewhat by the fact that rabbits appear to have significantly higher AM detection thresholds than humans (Carney et al. 2009).

To ensure that the rabbits in this study could readily perceive the S_{AM} (at least in the control conditions), a fully modulated signal was used in the neural recordings. This paradigm represents a departure from the stimulus parameters used in the human psychophysics, where the modulation depth was adaptively varied to estimate detection thresholds. However, a small sample of neurons tested with lower AM depths did not reveal a stronger tendency for neurons to show suppression when stimulus depths closer to human perceptual thresholds were used. Both of the hypothesized mechanisms for neural AM forward masking (suppression and contrast) are only meaningful when the S_{AM} elicits a change in a cell's firing in the control condition. Given the relatively high rate-based neural detection thresholds of rabbit IC neurons (and corresponding high behavioral thresholds), this issue becomes pertinent even with depths as high as 50% modulation. The

possibility cannot be completely dismissed that humans and rabbits may use or have access to different features of neural responses or that neural processing itself may be fundamentally different across the two species. Despite these caveats (which are always limitations of studies that attempt to relate physiology and behavior across species), the results presented here show clear qualitative agreement between changes in the contrast metric and the decay of psychophysical AM forward masking with increasing masker-signal delay, including a substantial residual effect of the AM masker at the longest delays used (~200 ms). Thus, it is likely that some general coding strategies are maintained across species and that the data provide relevant insights into the neural underpinnings of perceptual AM processing.

ACKNOWLEDGMENTS

REFERENCES

Bacon SP, Grantham DW (1989) Modulation masking: effects of modulation frequency, depth and phase. J Acoust Soc Am 85:2575–2580

Batra R (2006) Responses of neurons in the ventral nucleus of the lateral lemniscus to sinusoidally amplitude modulated tones. J Neurophysiol 96:2388–2398

Bendor D, Wang X (2007) Differential neural coding of acoustic flutter within primate auditory cortex. Nat Neurosci 10:763–771

Carney LH, Abrams KS, Koch K-J, Zilany MSA, Idrobo F (2009) Behavioral and physiological studies of amplitude-modulation detection. Meeting of Association for Research in Otolaryngology. Abstract 801

Chase SM, Young ED (2006) Spike-timing codes enhance the representation of multiple simultaneous sound-localization cues in the inferior colliculus. J Neurosci 26:3889–3898

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

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

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. J Acoust Soc Am 97:585–592

Ewert SD, Dau T (2000) Characterizing frequency selectivity for envelope fluctuations. J Acoust Soc Am 108:1181–1196

Ewert SD, Verhey JL, Dau T (2002) Spectro-temporal processing in the envelope-frequency domain. J Acoust Soc Am 112:2921–2931

Finlayson PG (1999) Post-stimulatory suppression, facilitation and tuning for delays shape responses of inferior colliculus neurons to sequential pure tones. Hear Res 131:177–194

Gilbert G, Bergeras I, Voillery D, Lorenzi C (2007) Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues. J Acoust Soc Am 122:1336

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

Gutschalk A, Micheyl C, Oxenham AJ (2008) The pulse-train auditory aftereffect and the perception of rapid amplitude modulations. J Acoust Soc Am 123:935–945

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

Joris PX, Yin TC (1992) Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am 91:215–232

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577

Krishna BS, Semple MN (2000) Auditory temporal processing: responses to sinusoidally amplitude-modulated tones in the inferior colliculus. J Neurophysiol 84:255–273

Langner G, Schreiner CE (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

Moore BCJ, Glasberg BR, Plack CJ, Biswas AK (1988) The shape of the ear's temporal window. J Acoust Soc Am 83:1102–1116

Moore BC, Fullgrabe C, Sek A (2009) Estimation of the center frequency of the highest modulation filter. J Acoust Soc Am 125:1075–1081

Nelson DA, Freyman RL (1987) Temporal resolution in sensorineural hearing-impaired listeners. J Acoust Soc Am 81:709–720

Nelson DA, Donaldson GS (2002) Psychophysical recovery from pulse-train forward masking in electric hearing. J Acoust Soc Am 112:2932–2947

Nelson PC, Carney LH (2007) Neural rate and timing cues for detection and discrimination of amplitude-modulated tones in the awake rabbit inferior colliculus. J Neurophysiol 97:522–539

Nelson PC, Smith ZM, Young ED (2009) Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking. J Neurosci 29:2553–2562

Noordhoek IM, Drullman R (1997) Effects of reducing temporal intensity modulations on sentence intelligibility. J Acoust Soc Am 101:498–502

Sakitt B (1973) Indices of discriminability. Nature 241:133–134

Strickland EA, Viemeister NF (1996) Cues for discrimination of envelopes. J Acoust Soc Am 99:3638–3646

Takahashi GA, Bacon SP (1992) Modulation detection, modulation masking, and speech understanding in noise in the elderly. J Speech Hear Res 35:1410–1421

Tan ML, Borst JG (2007) Comparison of responses of neurons in the mouse inferior colliculus to current injections, tones of different durations, and sinusoidal amplitude-modulated tones. J Neurophysiol 98:454–466

Wojtczak M, Viemeister NF (2005) Forward masking of amplitude modulation: basic characteristics. J Acoust Soc Am 118:3198–3210