



The rhythm of attention: Perceptual modulation via rhythmic entrainment is lowpass and attention mediated

Haleh Farahbod¹ · Kourosh Saberi¹ · Gregory Hickok^{1,2}

Published online: 19 July 2020
© The Psychonomic Society, Inc. 2020

Abstract

Modulation patterns are known to carry critical predictive cues to signal detection in complex acoustic environments. The current study investigated the persistence of masker modulation effects on postmodulation detection of probe signals. Hickok, Farahbod, and Saberi (*Psychological Science*, 26, 1006–1013, 2015) demonstrated that thresholds for a tone pulse in stationary noise follow a predictable periodic pattern when preceded by a 3-Hz amplitude modulated masker. They found entrainment of detection patterns to the modulation envelope lasting for approximately two cycles after termination of modulation. The current study extends these results to a wide range of modulation rates by mapping the temporal modulation transfer function for persistent modulatory effects. We found significant entrainment to modulation rates of 2 and 3 Hz, a weaker effect at 5 Hz, and no entrainment at higher rates (8 to 32 Hz). The effect seems critically dependent on attentional mechanisms, requiring temporal and level uncertainty of the probe signal. Our findings suggest that the persistence of modulatory effects on signal detection is lowpass in nature and attention based.

Keywords Periodicity · Entrainment · Rhythm

Rhythmic acoustic modulation occurs naturally in a large class of complex sounds, from speech and music to animal vocalizations and environmental sounds (Eddins & Bero, 2007; Elemans, Heeck, & Muller, 2008; Klump & Langemann, 1992; Lamminmaki, Parkkonen, & Hari, 2014; Peelle & Davis, 2012; Saberi & Hafter, 1995; ten Cate & Spierings, 2019). The human auditory cortex has evolved networks specialized for detecting the envelope spectrum of amplitude and frequency modulated signals (Barton, Venezia, Saberi, Hickok, & Brewer, 2012; Baumann et al., 2011; Hsieh, Fillmore, Rong, Hickok, & Saberi, 2012; Langner, Dinse, & Godde, 2009). These networks are most prominent in the core and belt regions of

the auditory cortex, are orthogonal to tonotopic gradients in auditory field maps, and have a lowpass characteristic with robust entrainment to modulation rates below 8 Hz (Barton et al., 2012; Joris, Schreiner, & Rees, 2004). Psychophysical findings are consistent with neurophysiological and neuroimaging results. Temporal modulation transfer functions (TMTFs), which measure modulation detection thresholds as a function of modulation rate, also have a lowpass characteristic for steady-state noise maskers with optimum detection at rates below 16 Hz and a shallow (3 dB/octave) roll-off above this rate (Eddins, 1993, 1999; Hsieh & Saberi, 2010; Scott & Humes, 1990; Morimoto et al., 2019).

One approach to measuring TMTFs is the probe-signal method in which the detection of a brief tonal pulse is measured as a function of its temporal position relative to the masker modulation phase. As modulation rate is increased, the detection of the probe becomes less dependent on modulator phase. This is largely due to a loss of phase-locking by auditory nerve fibers to the modulation envelope (Joris et al., 2004; Langner, 1992). The phase-dependent improvement in signal detection can be substantial. Scott and Humes (1990), for example, reported an approximately 40-dB improvement when a tone probe was presented at masker envelope minima relative to its

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13414-020-02095-y>) contains supplementary material, which is available to authorized users.

✉ Kourosh Saberi
saber@uci.edu

¹ Department of Cognitive Sciences, University of California, Irvine, CA, USA

² Department of Language Science, University of California, Irvine, CA, USA

maxima. Largest improvement was observed at the lowest modulation rate tested (2 Hz) with a steady decline in performance as modulation rate increased. Similar findings on detection of modulation signals have been reported using other methods for measuring TMTFs (Eddins, 1993, 1999; Morimoto et al., 2019; Scott & Humes, 1990).

More recently, a number of studies have investigated the predictive nature of modulating waveforms in signal detection. Luo and Poeppel (2007), for example, have reported that phase cues in low-frequency neural oscillations predict sentence intelligibility. Engel, Fries, and Singer (2001) and Giraud and Poeppel (2012) have found that rhythms in speech and other sounds provide predictive cues to the time of arrival of subsequent critical bits of information. Neural recordings in macaque auditory cortex have demonstrated a persistence of entrained oscillatory activity to a train of brief tonal pulses for several seconds after termination of acoustic stimulation (Lakatos et al., 2013). This neural persistence occurs only when the monkey is attending to the stimulus stream and not when the stream is ignored. Hickok, Farahbod, and Saberi (2015) have shown that psychophysical signal detection by human listeners is similarly affected by the phase of a 3-Hz modulating masker even after termination of modulation. Simon and Wallace (2017) used stimuli modeled after Hickok et al. to demonstrate a phase-dependent persistence of modulatory effects in EEG measures of a signal-detection task. Consistent with Hickok et al. (2015), they found that the EEG response to a (postmodulation) tone depended on the delay between the offset of modulation and onset of the tone in an antiphasic manner—that is, best signal detection occurred at the expected dip of the modulating masker (had the modulation continued). These findings collectively suggest a predictive role for modulating envelopes in auditory processing of attended sounds.

The current study extends the work of Hickok et al. (2015) by measuring TMTFs using a probe tone in stationary noise preceded by modulating maskers of different rates (2 to 32 Hz). The noise masker consisted of a modulating segment followed by a steady-state unmodulated segment during which the probe was presented. Thresholds were measured for the probe at various temporal positions after termination of masker modulation. We found a significant cyclic effect on signal detection for rates at or below 5 Hz that was entrained to the noise modulation envelope. Largest entrainment was observed at 2 and 3 Hz, with some residual effects at 5 Hz. No effects were observed at higher rates. In a second experiment, we found that persistence of the effects of modulation on signal detection requires signal uncertainty, suggesting a possible critical role for selective attention consistent with prior auditory neurophysiological findings (Lakatos et al., 2013).

Experiment 1

Method

Participants

Five normal-hearing (self-report) adults served as subjects in the 3, 5, 16, and 32 Hz conditions, and four normal-hearing adults in the 2 and 8 Hz conditions. Subjects were either undergraduate or graduate students at the University of California, Irvine (UCI), with the exception of one subject, who was a postdoc. All subjects were under 30 years of age. Subjects were unaware of the purpose of the experiment and were given instructions only on how to perform the task. Different subjects participated in different experimental conditions because of the extensive time it took to collect the large volume of data. Some subjects participated in multiple conditions. A total of 10 subjects ran in the various conditions of this study across three experiments. Detailed information about which subjects participated in which condition is available in the supplementary material uploaded to UCI's Data Repository (see Open Practices Statement at the end of the article). None of the authors served as a subject in this study.

Stimuli

Stimuli were generated using MATLAB software (The MathWorks, Natick, MA) on a Sony Lenovo T400 computer. Stimuli were presented at a rate of 44.1 kHz, through 16-bit digital-to-analog converters and Sennheiser headphones (eH 350) in a steel-walled acoustically isolated chamber (Industrial Acoustics Company). The masking stimulus was a broadband Gaussian noise burst with a nominal level of 70 dB (A weighted) measured using a 6-cc flat-plate coupler. The noise consisted of an initial modulated segment followed by a steady state (unmodulated) part. The total duration of the noise was 4 s for all modulation rates except for the 2-Hz condition, which had a duration of 4.5 s due to its long modulation period. The first ~3 seconds of the noise was sinusoidally amplitude modulated at a depth of 80%, terminating on the cosine phase of the first modulation cycle after 3 s (e.g., for a 5-Hz modulation rate, the modulating part was 3.1 s and the unmodulated part 0.9 s). Figure 1 shows an example stimulus. We employed six modulation rates of 2, 3, 5, 8, 16, and 32 Hz.¹ These rates include those typically associated with natural sounds, such as phonemes and syllables (<10 Hz), as well as some higher rates that allowed us to determine the upper boundary of potential modulatory effects. The signal to be detected was a 50-ms 1-kHz pure tone with a 5-ms rise-decay ramp for all modulation rates except for 32 Hz. Because the period of a 32-Hz modulator is brief (~31 ms), we

¹ Data for the 3-Hz condition are reproduced from Hickok et al. (2015).

set the duration of the tone signal for this condition to 5 ms with a 1-ms rise-decay time. This choice was arbitrary but seemed reasonable given the rapid modulation rate. The tone was centered at one of nine temporal positions during the unmodulated part of the noise burst. These temporal positions started at the offset of the modulation and were successively spaced at one-quarter of the modulation period (light blue circles in Fig. 1). Thus, the nine starting temporal position of the tonal signal covered two full cycles of the expected modulation waveform had the modulation continued during this period (i.e., yellow dashed curve).

Procedure

On each trial of a single-interval two-alternative forced-choice task, the subject was required to indicate (via a key press) whether a tonal signal was present during the unmodulated segment of the masking noise. Feedback was provided after each trial. The a priori probability of a signal occurring on a given trial was 0.5. When a tone was presented, its temporal position was selected randomly from one of nine delays, as shown in Fig. 1, and its level was selected randomly from one of five signal-to-noise ratios (SNR) covering a range of ~12 dB to allow measurement of psychometric functions. The 12-dB range was selected based on pilot work to produce a range of performance from near-chance to near-perfect detectability. Each run consisted of 100 trials and for most conditions, each subject completed a minimum of 27 runs per modulation rate, for a total of 77,300 trials across subjects and conditions.² This resulted in approximately 300 trials per delay per level per modulation rate. During each 2-hour session of the experiment, each subject completed approximately 8 to 10 runs. Each run lasted approximately 10 minutes. Subjects were given frequent breaks as needed and usually took a break after 2 to 3 runs in a session. Since all conditions were mixed within a run of 100 trials, performance was determined after completion of all runs for each subject by pooling data for a specific condition across all runs. All protocol were approved by the University of California, Irvine's, Institutional Review Board.

Results

For each subject and each modulation rate, psychometric functions were determined as a function of stimulus level

² The total number of trials collected for all experiments in the current study was 94,200, of which 77,300 were for the main experiment (six modulation rates) and 16,900 for the fixed-level experiment, which was restricted to only a single modulation rate of 3 Hz. Excluding the data of the 3-Hz condition, which was previously reported in Hickok et al. (2015), data from 86,400 new trials were collected in the current study. The number of trials in the main experiment of Hickok et al. was 11,600 from five subjects, and 14,400 for the entire study for an average of 28 runs per subject (this was previously stated as a minimum per subject). This included runs at signal-to-noise ratios, not shown. The 3-Hz data reproduced from Hickok et al. (was based on 7,800 trials from five subjects (900, 1000, 1900, 2000, 2000).

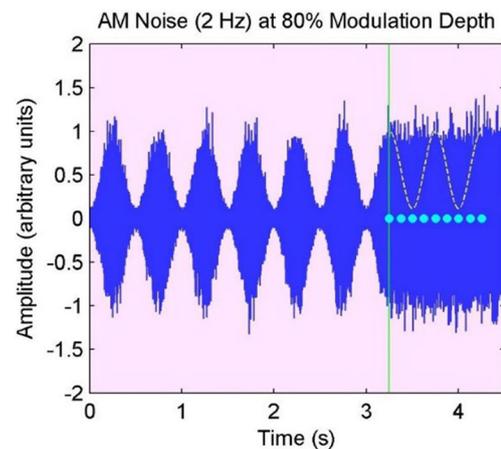


Fig. 1 Stimulus used in the current study. Modulating noise was followed by stationary noise. Tone probe was presented at one of nine temporal positions (light-blue circles). Vertical green line marks the end of masker modulation. Yellow sinusoidal curve shows the expected modulating envelope, had it continued. (Color figure online)

(collapsing across nine temporal positions). This resulted in a 5-point psychometric function at the five SNRs. Trials associated with the SNR on the psychometric function closest to the steepest point of its slope (between 0.7 and 0.9 proportion correct) were selected for further analysis, maximizing the likelihood of observing variations in performance (see Hickok et al., 2015).

The left panels of Fig. 2 show averaged proportion correct performance as a function of the delay between the offset of masker modulation and onset of tone signal (light blue circles in Fig. 1). Each panel shows performance for a different modulation rate from 2 to 32 Hz (top to bottom). Error bars are ± 1 standard error. Right panels of Fig. 2 show the same data as the left panels plotted as a function of the expected modulation phase at which the tone signal is presented. Note that the nine temporal positions at which the signal was presented cover exactly two full cycles of modulation, and hence the abscissa in the right panels cover 4π radians. Performance for the 2, 3, and 5-Hz conditions appear to be phase-locked and antiphasic to the expected modulation envelope, as is evident from the approximate M-shaped patterns (also see Supplemental Fig. S1). Performance for rates of 8 Hz or higher are inconsistent across the tone's temporal position and not phase-locked to the expected modulation cycle. Individual subject data are shown in Supplemental Figs. S2 and S3. Analysis of variance (ANOVA) (Greenhouse–Geisser correction for unequal variances³) showed a significant effect of the temporal position at which the signal was presented for modulation rates of 2 Hz, $F(8, 24) = 11.43$, $p = .005$, effect size $\eta^2 = 0.792$, observed power $\pi = 0.952$; 3 Hz, $F(8, 32) = 10.61$, $p = .004$, $\eta^2 = 0.726$,

³ Unequal variance is expected, given ceiling effects when proportion correct performance is near peak levels for some temporal positions, and near or below 0.7 for others (see top left panel of Supplemental Fig. S2).

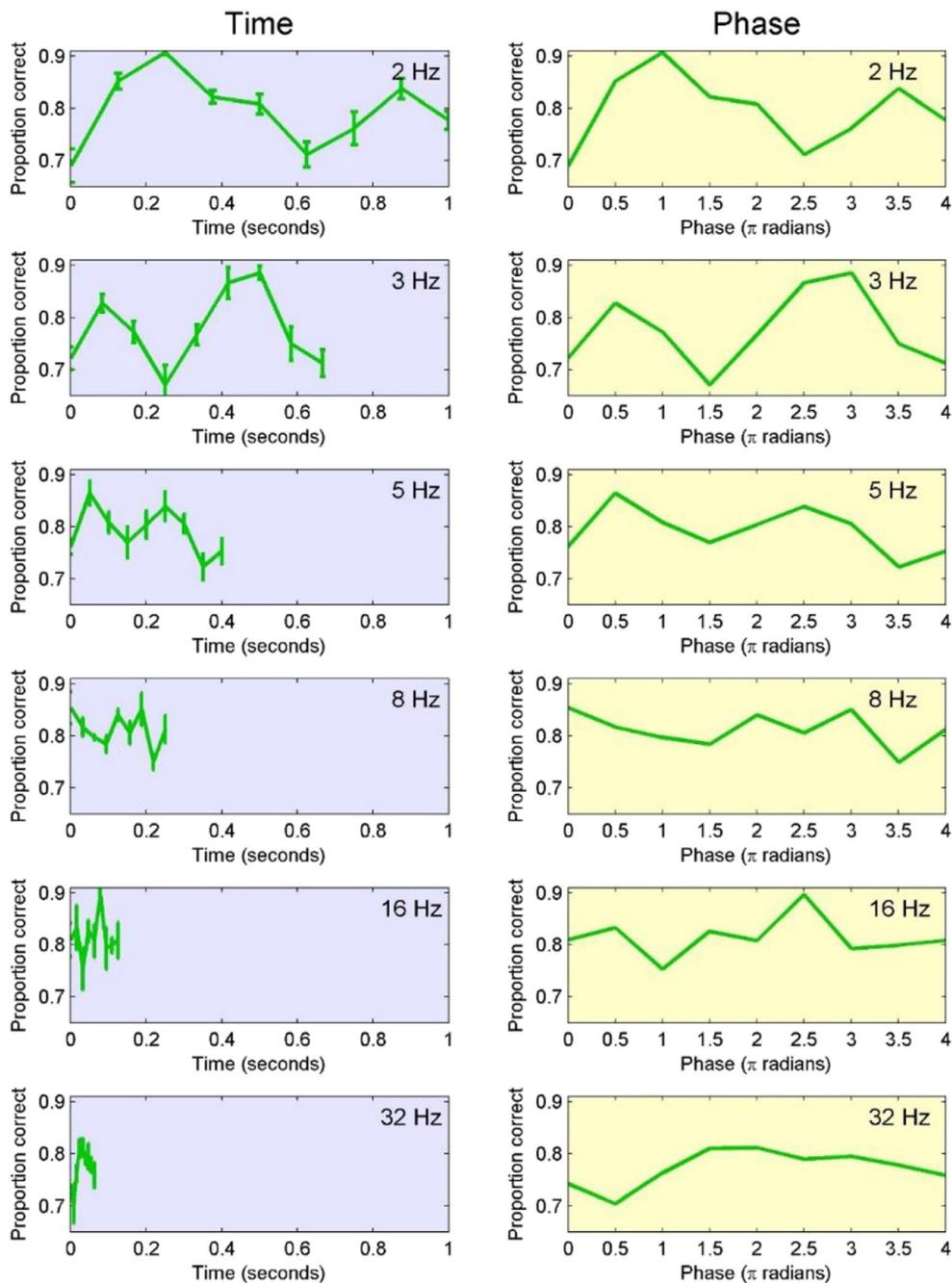


Fig. 2 Left panels show tone detectability as a function of time after offset of noise modulation. Each panel shows data from one modulation rate (top to bottom: 2 to 32 Hz). The abscissa represents time after offset of masker modulation (at its cosine phase). Error bars are ± 1 standard

error. Right panels show the same data plotted as a function of the expected modulation phase at which the tone signal is presented (two cycles: 0 to 4π)

$\pi = 0.944$; and 5 Hz, $F(8, 32) = 4.87, p = .041, \eta^2 = 0.549, \pi = 0.631$, but no significant effects for 8 Hz, $F(8, 24) = 2.50, p = .182, \eta^2 = 0.455, \pi = 0.275$; 16 Hz, $F(8, 32) = 2.69, p = .092, \eta^2 = 0.402, \pi = 0.512$, or 32 Hz, $F(8, 32) = 3.31, p = .078, \eta^2 = 0.453, \pi = 0.509$. Bayes factor analysis (BIC approximation) showed moderate evidence for the null hypothesis at the 16-Hz modulation rate $BF_{01} = 6.51$, and strong evidence for the null at 32 Hz, $BF_{01} = 15.16$. For the three statistically

significant modulation rates (2, 3, 5 Hz), d' values are shown in Supplementary Fig. S1 for comparison, measured as $z(\text{hits}) - z(\text{false alarms})$ individually for each subject and signal delay (Macmillan & Creelman, 2004).

We also conducted additional analyses to get a better sense of the temporal periodicity effects observed in the data. Several measures were examined. First, we measured the autocorrelation function for the data of Fig. 2, as well as the

cross-correlation function of these data with a single cycle of a sinusoid at the expected modulation rate. We found peaks in the cross-correlation function at the expected rates for rates at or below 5 Hz and no discernable effects above 5 Hz. However, both of these measures (cross-correlation and autocorrelation functions) were relatively noisy and largely dominated by the peak at zero delay given the very brief duration of the “waveforms” (2 cycles). Second, we examined the Fourier spectrum of the functions shown in Fig. 2 to determine the magnitude of spectral peaks at the expected modulation rate relative to the level of background noise at other frequencies. This approach was less informative as the duration of functions shown in Fig. 2 are exactly two cycles of modulation, producing spectral peaks at half the expected modulation rate (i.e., $1/\text{duration}$) whether or not there is any periodicity. This, in turn, may produce spectral peaks at the first harmonic of $1/\text{duration}$, which is equal to the expected modulation rate (i.e., a false positive). Overall, our analyses using a number of approaches show that the strongest entrainment effects are at the lowest modulation rates tested.

One other interesting observation is worth noting. There appears to be a shift in the minima (and maxima) of the sustained periodicity in signal detection with increasing modulation rate. This can be observed in the right panels of Fig. 2 in which performance is plotted as a function of modulation phase. The top right panel (2 Hz) shows a minimum at a phase angle slightly higher than 2π radians, whereas the middle panel (3 Hz) shows this dip at slightly below 2π . This phase shift may also suggest a possible slight shift in the frequency of the detection functions relative to the referent (stimulus) modulation rates. This can be seen in the same three panels where the 2 Hz data seem to be associated with a frequency slightly lower than 2 Hz (we estimate this at 1.4 Hz), and the 3 and 5 Hz data seem to be associated with a frequency slightly greater than their own referent rates. A similar and possibly related effect is observed in the data of Scott and Humes (1990). They show that increasing the masker modulation rate results in a shift in the minima of the functions that relate signal-detection thresholds to the phase of the modulating masker (see Fig. 2 of Scott & Humes, 1990).

Experiment 2

Signal uncertainty and selective attention

Signal uncertainty has been shown to enhance selective attention in signal detection (Dieterich, Endrass, & Kathmann, 2016). As uncertainty increases (and predictability decreases), the system’s limited attentional resources are allocated to monitoring an increasing number of sensory channels. The use of an “attentional filter” under conditions of uncertainty has been shown to improve performance in a number of basic

auditory tasks, including the detection of simple or complex tones of uncertain frequency (Dai, Scharf, & Buus, 1991; Hafer & Saberi, 2001; Schlauch & Hafer, 1991; Wright & Fitzgerald, 2017) uncertain duration (Dai & Wright, 1995), or uncertain time of occurrence (Bourbon, Hafer, & Evans, 1966). To our knowledge, no study has examined the effects of stimulus level uncertainty on selective attention. Neuroimaging studies (Bilecen, Seifritz, Scheffler, Henning, & Schulte 2002; Schreiner & Malone, 2015) have interestingly identified “amplitopic” maps in the human auditory cortex (transverse temporal gyrus) that appear spatially arranged along stimulus-intensity gradients. These intensity maps can potentially serve as substrates for focusing attentional filters on specific stimulus levels.

Our rationale for considering the role of uncertainty and attention in persistence of modulatory effects was as follows. If amplitude modulation during early parts of the masking noise modulates attention, and if signal uncertainty during the steady state portions of the noise enhances use of selective attention during periods of minimum noise, then the persistence of a brief attentional cadence after termination of masker modulation may yield results consistent with the observed antiphase pattern of signal detection. If level uncertainty is removed, the role of selective attention may diminish, resulting in a flat detection function.

We should note that temporal uncertainty obviously remains after removal of level uncertainty. Prior work has demonstrated a strong influence of temporal uncertainty on auditory signal detection (Bonino & Leibold, 2008; Bonino, Leibold, & Buss, 2013; Bourbon et al., 1966), and we suspect that removal of temporal uncertainty would also likely diminish the periodicity of postmodulation signal detection. However, for two reasons, we focused on level uncertainty. First, virtually no studies have examined the effects of level uncertainty on auditory signal detection, and we therefore thought that this would be an interesting stimulus dimension to investigate, especially given discoveries of amplitopic maps in auditory cortex (Bilecen et al., 2002; Schreiner & Malone, 2015). Second, pilot tests during early stages of the experiment unexpectedly showed that modulatory effects diminish when using fixed-level stimuli, and hence, the current study was designed to further investigate those early observations.

Method

All stimulus parameters, experimental design, and equipment were the same as those used in Experiment 1, except for the following. The signal level was fixed within a run instead of randomly selected from a set of five levels. A single masker modulation rate of 3 Hz was used, and data were collected at four of the five different SNRs (fixed within a run). Three subjects who had participated in Experiment 1 also

participated in this study, with each subject completing approximately 40 runs of 100 trials each.

Results

Figure 3a (top left) shows results of this experiment for two of the fixed signal levels, with the upper curve showing data for a tone that was 3.5 dB higher than that of the lower curve. The other two levels investigated produced ceiling-level performance and were therefore not informative. Each dotted trace shows data from one listener, and the bold solid line shows their average. No modulatory effect on signal detection is observed. This is contrary to the pattern of performance shown in Fig. 3b for the *same* three subjects in the multilevel condition where a clear effect of masker modulation is observed. The data of Fig. 3b are pooled from those trials on which the signal level was 3.5 dB, corresponding to the upper curve in Fig. 3a. Figure 3c shows the difference between the mean performance in the random (multilevel) and fixed-level conditions at the 3.5-dB signal-level condition. This is the

difference between the solid line in panel b and the top solid line in panel a. A two-way (2×9) repeated-measures ANOVA on the data of the left panels (at 3.5 dB) showed a significant effect of condition (fixed vs. multi), $F(1, 2) = 20.77, p = .045$, a near-significant effect of signal delay, $F(8, 16) = 7.17, p = .075$, and a near-significant interaction effect between condition and delay, $F(8, 16) = 5.78, p = .074$. The nonsignificance of the “delay” and interaction effects is not surprising, as the relatively flat functions of the fixed-level conditions diminish the overall modulatory effects when data are collapsed across fixed and multilevel conditions. While there does seem to be some trend toward lower performance at the shortest and longest delay in the fixed-level condition, there is no cyclic pattern as seen in the random-level conditions. A one-way ANOVA on only the fixed-level data showed no significant effect of signal delay on performance, $F(8, 16) = 3.33, p = .157$, and Bayes factor analysis on the same data provided further evidence in support of the null hypothesis that signal delay has no significant effect on performance in the fixed-level condition, $BF_{01} = 2.10$.

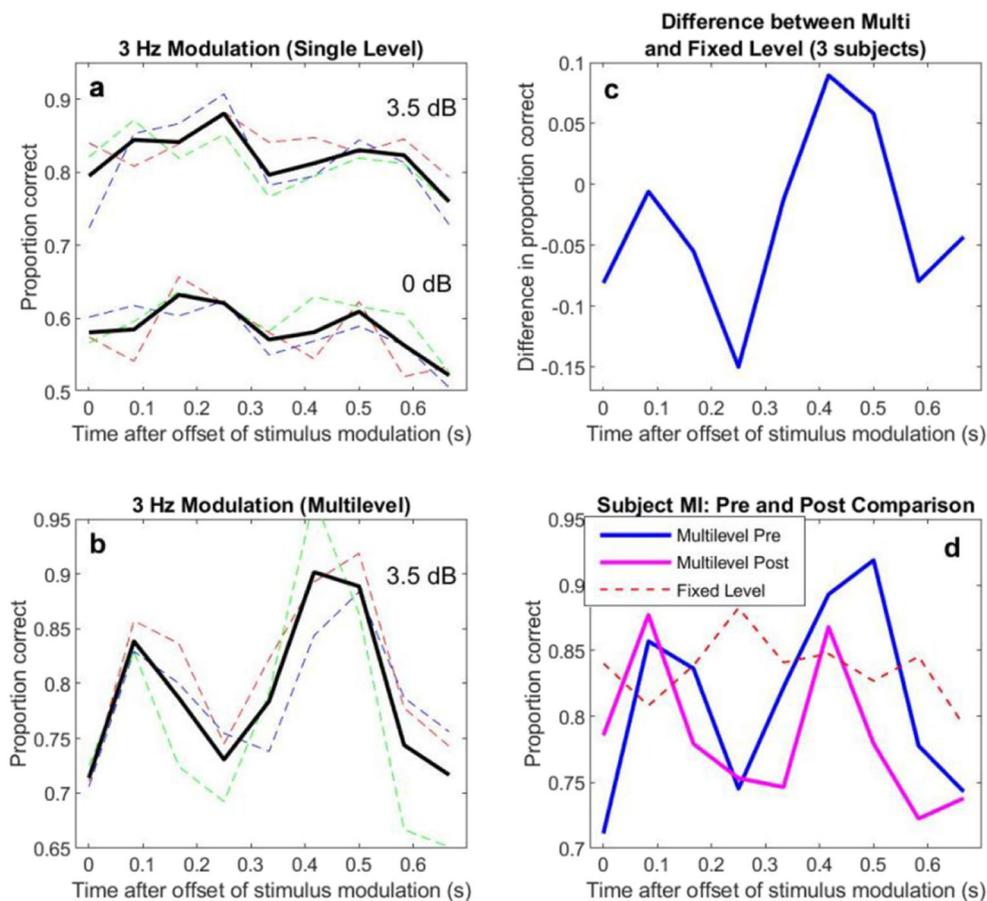


Fig. 3 **a** Performance for three subjects when signal level was fixed within a run (no uncertainty). Data are shown for two different signal levels. The heavy solid lines show their mean. **b** Data from the same three subjects when signal level was randomized from trial to trial across five levels within a run (multilevel uncertainty). Note that data

from only one signal level (3.5 dB) out of five is shown even though signal level was roved on every trial. **c** Difference between the random and fixed-level conditions (at 3.5 dB signal level). **d** Performance for one subject (MI) in multilevel condition before and after being tested at the fixed-level condition (see text for details)

For one subject, we repeated the original experiment with level uncertainty *after* the subject had completed the single-level condition. Thus, this subject completed the original multilevel experiment, followed by the fixed-level experiment, and then followed by a replication of the multilevel experiment. Results, plotted in Fig. 3d, show that reintroduction of level uncertainty restores the modulatory pattern of signal detection for this subject.

One other interesting observation is worth noting. The average proportion-correct performance level, pooled across all nine signal delays, is nearly identical for the fixed and random level conditions (0.82 vs. 0.79, respectively), $t(25) = 1.95$, ns). These two proportion correct values are the average of the top bold curve in Fig. 3a and the average of the bold curve in Fig. 3b. However, the peaks and dips of the difference function (Fig. 3c) are significantly different from each other, $t(2) = 5.04$, $p = .037$. This suggests that under conditions of level uncertainty, signal detection is enhanced at the expected dips of the masker modulation envelope and degraded at the expected peaks. In other words, best performance across delays occurs in the level uncertainty condition (not fixed level), possibly as a result of focal attention at the expected dips.

Finally, while it is not unusual for auditory experiments to employ three subjects per condition (Hsieh, Petrosyan, Goncalves, Hickok, & Saberi, 2011; Hsieh & Saberi, 2010), and while several statistical tests converge in support of the main conclusions of this experiment, one should nonetheless exercise some caution in interpreting these results due to the comparatively smaller number of subjects. Specific support, however, for a relatively robust effect includes (1) a statistically significant difference between fixed-level and random-level conditions; (2) an absence of an effect, as predicted, in the *fixed-level* condition that was additionally supported by Bayes factor analysis; and (3) completion of over 16,000 trials by the three subjects in a within-subject's design, which, even for psychophysical experiments, is a relatively large number.

Discussion

Hickok et al. (2015) reported that low-frequency rhythmic stimulation can induce an oscillation in signal detection even after the driving stimulus has stopped oscillating. The present study extended these findings to a wide range of modulation rates to better characterize the nature of this persistent postmodulatory effect. Prior work has shown that sensitivity to amplitude-modulation detection in broadband noise maskers generally has a lowpass shape with greatest sensitivity at rates below 8 Hz. We found a similar lowpass effect in our study, but with a significantly more pronounced decline in performance with increasing modulation rate.

Traditional TMTFs have shown a shallow roll-off with a slope of approximately 3 dB per octave as a function of modulation rate. Significant effects of modulation have been

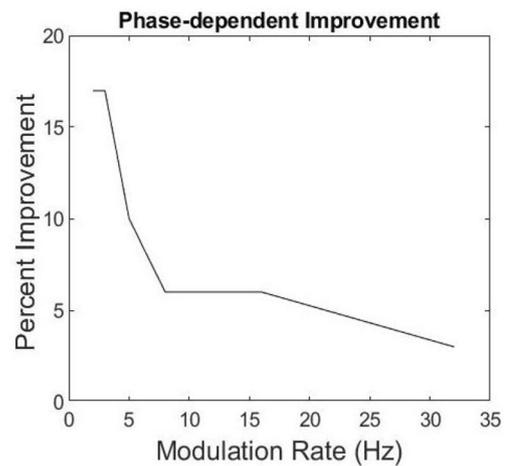


Fig. 4 Estimated change in signal-detection performance as a function of modulation rate. Sinusoidal functions were fitted to the data of Fig. 2, and improvement in performance was measured as the difference between peaks and troughs of the fitted functions

reported even at rates as high as 32 Hz. This is in contrast to our finding that the persistence of entrainment to modulation precipitously decays for rates above 3 Hz.

Figure 4 shows estimated change in signal detection thresholds from the current study as a function of masker modulation rate. To obtain these estimates, we fit sinusoidal functions to the data of Fig. 2 and measured difference in performance between peaks and troughs of the fitted functions (see Supplemental Fig. S4). Change in performance, which is a measure of the degree to which performance has entrained to the modulation envelope, clearly drops sharply for rates above 3 to 5 Hz. While some improvement may be observed for higher rates, these are likely due to noise in performance, as no consistent modulation-phase dependence was observed across individual subjects at these higher rates. This is confirmed by the statistically nonsignificant effects of probe delay at these rates. The shape of the function shown in Fig. 4 is not surprising, given the well-known loss of cortical phase locking with increasing modulation rate. Virtually no phase locking remains above a modulation rate of about 20 Hz (Joris et al., 2004). This finding is also consistent with the neurophysiological work of Lakatos et al. (2013) who demonstrated sustained *postmodulation* phase locking to a driving modulatory auditory stimulus in monkey cortical neurons for rates up to 6 Hz, but not for a rate of 12 Hz.

Predictions from a modulation filterbank model

To gain better insight into the signal processing dynamics involved in the observed modulatory effects, we conducted a more detailed analysis of how the auditory system processes modulating waveforms of the type used in our experiments. The lowpass transfer function typically observed for modulation detection may be more accurately characterized as

resulting from the output of a bank of bandpass modulation filters orthogonal to the conventional tonotopically organized peripheral (critical band) filters (Dau, Püschel, & Kohlrausch, 1996). There is neurophysiological and neuroimaging support for existence of such modulation filters at the cortical level, and psychophysical modeling has provided compelling support for such filters (Barton et al., 2012; Dau, Kollmeier, & Kohlrausch, 1997; Hsieh & Saberi, 2010; Sek & Moore, 2003; Xiang, Poeppel, & Simon, 2013). To this end, we processed the 3-Hz modulated noise stimulus used in our experiment through a model of the auditory periphery followed by a modulation filterbank. The initial peripheral processing consisted of a GammaTone filterbank with 50 filters whose center frequencies (CFs) were logarithmically spaced from 500 to 2,000 Hz (Holdsworth, Nimmo-Smith, Patterson, & Rice, 1988; Slaney, 1998) followed by an inner hair-cell model (Meddis, Hewitt, & Shackleton, 1990; Slaney, 1998). The output of this model is shown in Fig. 5. Left panels show a single stimulus (masker plus 1 kHz tonal signal) and right panels show average of 25,000 stimuli (masker only). To facilitate visual inspection, the 1-kHz tone was set to the highest SNR used in the experiment at a temporal position corresponding to the first expected dip in the modulating envelope had the modulation continued (third temporal position in Fig. 1). Bottom panels show the same model output, except that filter outputs were integrated across frequency channels to better show overall changes in amplitude of the noise envelope. Note that the masking noise causes greater activity in the outputs of the higher frequency filters (top-right panel) because of the increasing bandwidth of peripheral filters with increasing center frequency. Second, and more importantly,

note the monotonic decline in the average amplitude of the filter outputs immediately after termination of modulation (bottom-right panel). This decline may partially explain why detection of a tone signal immediately after termination of modulation (first delay point) is poor, but cannot explain the modulatory nature of performance (M-shaped function) that persists for two modulation cycles after termination of the driving noise modulator.

The model output from the peripheral filter centered at 1 kHz was then extracted and processed through a modulation filterbank. The reason for this additional step was the concern that perhaps modulation filters would “ring” in an oscillatory manner that could predict the M-shaped functions observed in the behavioral data. The model comprised five modulation filters with a Q of 1 and resonant frequencies of 3, 6, 12, 24, and 48 Hz (Dau et al., 1996; Hsieh & Saberi, 2010; Sek & Moore, 2003). Top panel of Fig. 6 shows the output of the modulation filterbank for the 3-Hz noise masker (no signal). Bottom panel shows outputs of the five individual filters.

Note that, as expected, the greatest activity is observed at 3 Hz and that after termination of masker modulation, the output of the modulation filters decline to zero for most filters. For the modulation filter centered at 3-Hz (blue line in the bottom panel of Fig. 6), this activity monotonically declines, persists for several hundred milliseconds, but not with an oscillatory profile. Thus, while the output activity of modulation filters may explain part of the observed behavioral data (i.e., poorer performance at the transition point between the modulating and steady state sections of the noise masker), we do not see a sinusoidal ringing of these filters after termination of

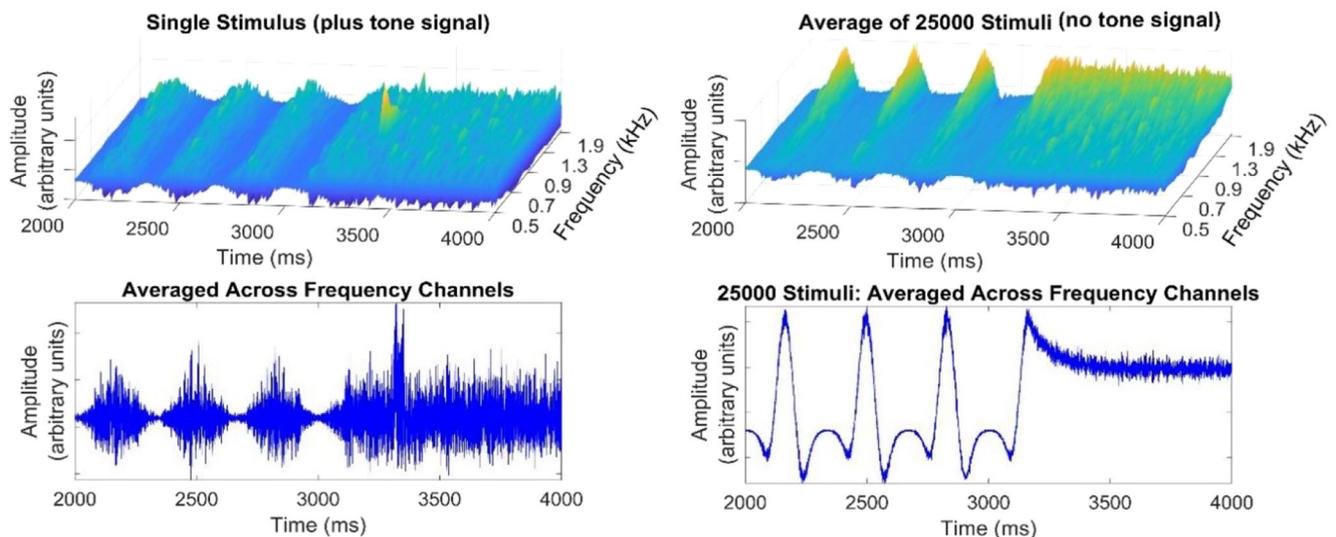


Fig. 5 Left panels show the output of a model of the auditory periphery in response to the type of stimuli used in our experiment (3-Hz masker condition plus a 1-kHz tonal signal). Right panels show the averaged model output in response to 25,000 masker samples (no signal). Bottom

panels show model outputs integrated across frequency channels. Note the decline in the masker noise amplitude envelope immediately after termination of the 3-Hz modulation (bottom-right panel). See text for details

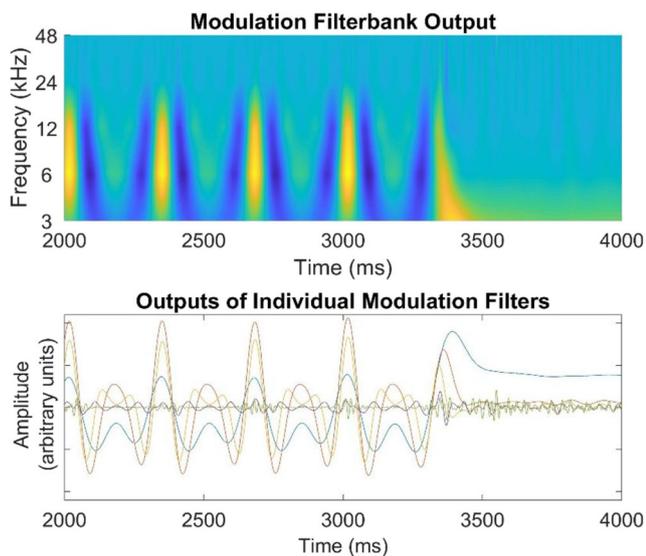


Fig. 6 Output of a modulation filterbank model in response to the 3-Hz noise masker used in the current study. The output of the peripheral auditory filter at 1 kHz (from Fig. 5) was used as input to the modulation filterbank. Top panel shows the model's response as a function of time and filter center frequency. Bottom panel shows the response of each individual filter separately. The blue line in the bottom panel represents the response of the modulation filter centered at 3 Hz. (Color figure online)

modulation in a way that could explain the M-shaped pattern of performance seen in Fig. 2.

Monte Carlo simulations

We combined the model output with a putative attention modulation function to predict psychophysical performance using Monte Carlo simulations. The simulations were motivated by two considerations. First, how does the decline in the amplitude of the noise envelope after termination of the driving modulator (at the model's output) affect predicted performance. Second, if modulation of attention induces a change in effective SNR, how much change in SNR predicts the approximately 20% change in percentage-correct performance observed in the M-shaped functions of Fig. 2.

The Monte Carlo simulation consisted of 1,000 runs of 100 trials each using the same stimuli as in the 3-Hz condition of Experiment 1, with a tone signal level that generated the data in the second row of Fig. 2. Each signal-plus-noise stimulus was independently generated on each trial and filtered through the peripheral filterbank model described earlier. The peak value of the model output during the steady state part of the masker (after termination of modulation) was then compared with a criterion. If the output amplitude at any point exceeded this criterion, a “signal present” response was recorded, and otherwise a “no signal” was recorded. The value of the criterion was a free parameter of the model and set to produce an average performance (across all delays) that matched that of

the actual average performance observed in the psychophysical experiments (approximately 80% correct). Modulation of attention was simulated by weighting the tone signal amplitude with a sinusoidal function antiphase to the masker modulation envelope. Obviously, a sinusoidal weighting function will likely result in a sinusoidal pattern of performance. However, as noted earlier, our motivation for the simulation was twofold. First, to determine the degree to which a postmodulation decline in the envelope amplitude (as seen in Figs. 5 and 6) affects the pattern of simulated performance, and second, to determine how much change in SNR (from the sinusoidal attention function) would produce simulation results that matched the range of change in proportion-correct performance observed in the behavioral data. The weighting function was $W(\tau) = 1 + m [\sin(2\pi f_m \tau - \pi/2)]$, where τ is the temporal position of the tone signal as shown in Fig. 1 (one of nine values), f_m is modulation frequency (3 Hz in the current simulation), and m is the *signal* modulation depth (i.e., the amplitude by which the putative attentional mechanism modulates the “effective” signal level). The amplitude parameter (m) was the second free parameter of the model. No internal noise was added to the process since the external (masking) noise had a large enough variance to generate the probabilistic range of performance observed in the psychophysical experiments.

Simulation results, shown in Fig. 7, suggest that it takes very little “effective” signal modulation ($m = 0.18$, -14.9 dB modulation depth) to generate the approximately 20% range of change in performance observed in the data. If attention is in fact modulated by the masker's driving modulator, and if this attentional modulation briefly persists during the steady state portion of the masking noise, then a $\sim 20\%$ change in performance can be induced with as little as 0.18 perturbation of the amplitude of the tonal signal. The simulation results

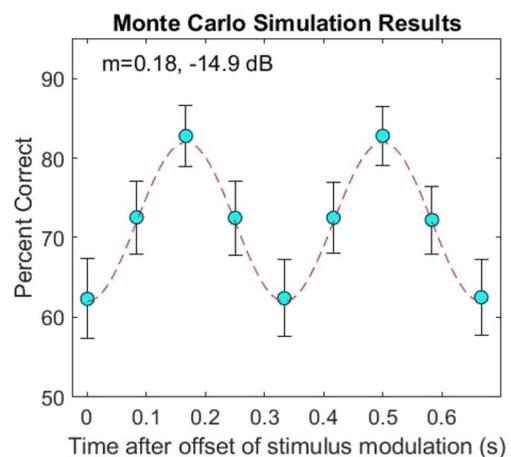


Fig. 7 Results of the Monte Carlo simulation (3-Hz condition) based on 1,000 runs of 100 trials each. The approximately 20% range of change in performance observed in the data can be generated with very little “effective” signal modulation ($m = 0.18$, -14.9 dB modulation depth)

also show virtually no effect of the decline in noise amplitude envelope immediately after termination of the driving masker modulation (bottom-right panel of Fig. 5). While such a decline in noise masker envelope may be a real effect of how peripheral filters operate, these processes will likely similarly affect the amplitude of the tonal signal, leaving SNR relatively unchanged. In addition, the variance of amplitude of a masker sample on a given trial perhaps is too large relative to any single-trial decline in postmodulation amplitude envelope to affect performance. In other words, the decline in postmodulation amplitude envelope becomes evident after averaging 25,000 noise masker samples and may be relatively inconsequential on a trial-by-trial basis (compare the bottom left and right panels of Fig. 5).

Comparison to findings from vision research

Substantial evidence from vision research suggests that periodic modulatory effects are a natural feature of cortical activity that could be linked to attentional mechanisms (Buzsáki, 2006; VanRullen & Macdonald, 2012). Psychophysical studies suggest that selective attention samples visual features (e.g., location) rhythmically at a rate of about 8 Hz. As demand for attentional resources is increased—for example, by having to simultaneously monitor two visual locations or features—the system allocates resources proportionately and samples each at around 4 Hz (Fiebelkorn, Saalman, & Kastner, 2013; Landau & Fries, 2012; Re, Inbar, Richter, & Landau, 2019).

The phase of oscillatory patterns of psychophysical performance can be reset both in visual (Landau & Fries 2012) and auditory tasks (Ho, Leung, Burr, Alais, & Morrone, 2017). For example, a brief visual flash at one location can reset this phase in a visual change-detection task involving two target locations with an antiphase pattern of performance associated with the two loci. Similarly, in a pitch-discrimination task, the onset of steady state noise triggers a reset of the phase of an oscillatory process that cyclically affects the discrimination of the pitch of a brief tone probe (Ho et al., 2017). The cyclic pattern of psychophysical performance (as a function of tone delay) is ear dependent, with the left and right ears producing an antiphase pattern. The authors speculated that the antiphase nature of this pattern across ears may possibly explain why some auditory studies have failed to show periodic patterns in psychophysical performance. They suggest that because most of these studies use diotic signals (presented simultaneously to both ears), any potential modulatory effect would be canceled when summed across ears. Although we used diotic signals in our current study, our design differs from Ho et al. in two major ways. First, we used a tone-in-noise detection task whereas Ho et al. used a pitch-discrimination task. It is unclear whether the antiphase processes across ears observed by Ho et al. may extend to a tone-

in-noise detection task. Second, our steady state masking stimuli were preceded by amplitude modulated noise, which may have synchronized the phase of oscillatory attentional effects at the two ears. Phase resets have also been shown across modalities, with an auditory stimulus triggering a reset of the phase of oscillatory visual behavioral and cortical activity patterns (Mercier et al., 2013; Romei, Gross, & Thut, 2012), and conversely, a visual stimulus resetting the phase of oscillatory patterns associated with the auditory system (Thorne, De Vos, Campos Viola, & Debener, 2011).

Finally, in another interesting vision study, Spaak, de Lange, and Jensen (2014) used a stimulus design more similar to ours in investigating both psychophysical performance and cortical oscillatory activity patterns after termination of a periodic visual stimulus. They reported that cortical activity patterns outlasted the stimulus in a periodic manner for several cycles. Furthermore, they found that psychophysical performance in a visual signal-detection task mirrored the cyclic pattern of poststimulus cortical activity. This finding is consistent with auditory neurophysiological findings in macaque monkeys, as well as with our current psychophysical results that thresholds may show a cyclic pattern after termination of the driving stimulus. Our findings, together with those of other auditory and visual experiments, suggest that attention-based rhythmic modulation of signal detection may be a universal feature of perceptual systems and not modality specific.

Conclusions

In our original paper, we speculated that the persistence of modulatory effects may possibly be a bottom-up process (Hickok et al., 2015). This idea was partly a result of the antiphase shape of the sustained modulatory effect with respect to the driving modulation envelope. Our reasoning was that if listeners had in fact attended to the peak of the amplitude-modulated noise to predict stimulus arrival, one would expect best performance at the peak of the expected modulation cycle contrary to what we had observed. Our current results, however, suggest that attention likely does play a critical role in observing sustained entrainment. When signal level uncertainty is eliminated from the experimental design, persistence of modulatory effects disappears, and when it is reintroduced, the original effect is recovered. How, then, can one interpret the antiphase shape of entrainment if it is actually driven by attentional processes? Perhaps attention cues the subject to listen at points in time when the SNR is at its maximum (listening in the dip strategy). Listeners have been shown to take advantage of “glimpses” in the dips of modulated noise to detect a signal by directing attention to the portions of a signal with the most favorable SNR (Festen & Plomp, 1990; Hopkins & Moore, 2009; Peters, Moore, & Baer, 1998). Neurophysiological findings also support a role for attention in persistence of modulatory effects. For

example, animals that ignore the modulating auditory stimulus do not show sustained neural entrainment after termination of the driving modulator (Lakatos et al., 2013). EEG measurements during signal-detection tasks in humans also implicate potential attentional mechanisms in postmodulatory persistence (Simon & Wallace, 2017).

In summary, we have found that when an amplitude modulated noise is followed immediately by a stationary noise masker, the detectability of a probe tone in the stationary noise follows a cyclic pattern as if the modulating masker had continued. This effect is observed for approximately two cycles of the driving modulation envelope. The strongest entrainment occurs at the lowest modulation rates of 2 and 3 Hz, with some residual effects at 5 Hz, and no effect at higher rates. The effect also seems to be at least partly mediated by attentional mechanisms and signal uncertainty. An interesting question that remains is whether entrainment effects can generalize to other types of biologically relevant auditory signals. One such signal is the modulation envelope itself. If the signal to be detected is a single-cycle of AM with a shallow (near threshold) depth, does its detection depend on its onset phase relative to the terminating phase of the driving modulator? We suspect that AM detection would be worse when its starting phase matches the ending phase of the driving modulator masker, and best when they are antiphasic. Furthermore, we speculate that a mismatch between the rates of the AM signal and that of the driving modulator may improve AM detection. Should we find entrainment for AM (or FM) signals comparable to those of the current study, it would suggest that persistence of modulatory effects can generalize to a wide class of basic auditory signals including potentially speech and music.

Acknowledgements Work supported by NIH R01DC009659. We thank two anonymous reviewers for their helpful suggestions.

Open practices statement The raw data for all experiments as well as other supplementary information are available at the University of California, Irvine's Data Repository (Dryad): <https://doi.org/10.7280/D1WD5D>

References

- Barton, B., Venezia, J. H., Saberi, K., Hickok, G., & Brewer, A. (2012). Orthogonal acoustic dimensions define auditory field maps in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 20738–20743. doi:<https://doi.org/10.1073/pnas.1213381109>
- Baumann, S., Griffiths, T. D., Sun, L., Petkov, C. I., Thiele, A., & Rees, A. (2011). Orthogonal representation of sound dimensions in the primate midbrain. *Nature Neuroscience*, *14*, 423–425. doi:<https://doi.org/10.1038/nn.2771>
- Bilecen, D., Seifritz, E., Scheffler, K., Henning, J., & Schulte, A. C. (2002). Amplitude of the human auditory cortex: an fMRI study. *NeuroImage*, *17*, 710–718. doi:<https://doi.org/10.1006/nimg.2002.1133>
- Bonino, A. Y., & Leibold, L. J. (2008). The effect of signal-temporal uncertainty on detection in bursts of noise or a random-frequency complex. *Journal of the Acoustical Society of America*, *124*, EL321–EL327. doi:<https://doi.org/10.1121/1.2993745>
- Bonino, A. Y., Leibold, L. J., & Buss, E. (2013). Effect of signal-temporal uncertainty in children and adults: Tone detection in noise or a random frequency masker. *Journal of the Acoustical Society of America*, *134*, 4446–4457. doi:<https://doi.org/10.1121/1.4828828>
- Bourbon, W. T., Hafter, E. R., & Evans, T. R. (1966). Frequency and time uncertainty in auditory detection. *Journal of the Acoustical Society of America*, *39*, 1247. doi:<https://doi.org/10.1121/1.1942836>
- Buzsáki, G. (2006). *Rhythms of the brain*. Oxford, England: Oxford University Press. doi:<https://doi.org/10.1093/acprof:oso/9780195301069.001.0001>
- Dai, H., & Wright, B. A. (1995). Detecting signals of unexpected or uncertain durations. *Journal of the Acoustical Society of America*, *98*, 798–806. doi:<https://doi.org/10.1121/1.413572>
- Dai, H., Scharf, B., & Buus, S. (1991). Effective attenuation of signals in noise under focused attention. *Journal of the Acoustical Society of America*, *89*, 2837–2842. doi:<https://doi.org/10.1121/1.400721>
- Dau, T., Püschel, D., & Kohlrausch, A., (1996). A quantitative model of the “effective” signal processing in the auditory system: I. Model structure. *Journal of the Acoustical Society of America*, *99*, 3615–3622. doi:<https://doi.org/10.1121/1.414959>
- Dau, T., Kollmeier, B., & Kohlrausch, A., (1997). Modeling auditory processing of amplitude modulation: II. Spectral and temporal integration. *Journal of the Acoustical Society of America*, *102*, 2906–2919. doi:<https://doi.org/10.1121/1.420345>
- Dieterich, R., Endrass, T., & Kathmann, N. (2016). Uncertainty is associated with increased selective attention and sustained stimulus processing. *Cognitive, Affective, & Behavioral Neuroscience*, *16*, 447–456. doi:<https://doi.org/10.3758/s13415-016-0405-8>
- Eddins, D. A. (1993). Amplitude modulation detection of narrow-band noise—Effects of absolute bandwidth and frequency region. *Journal of the Acoustical Society of America*, *93*, 470–479. doi:<https://doi.org/10.1121/1.405627>
- Eddins, D. A. (1999). Amplitude-modulation detection at low- and high-audio frequencies. *Journal of the Acoustical Society of America*, *105*, 829–837. doi:<https://doi.org/10.1121/1.426272>
- Eddins, D. A., & Bero, E. M. (2007). Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region. *Journal of the Acoustical Society of America*, *121*, 363–372. doi:<https://doi.org/10.1121/1.2382347>
- Elemans, C. P. H., Heeck, K., & Muller, M. (2008). Spectrogram analysis of animal sound production. *Bioacoustics: The international Journal of Animal Sound and its Recording*, *18*, 183–212. doi:<https://doi.org/10.1080/09524622.2008.9753599>
- Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, *2*, 704–716. doi:<https://doi.org/10.1038/35094565>
- Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *Journal of the Acoustical Society of America*, *88*, 1725–1736. doi:<https://doi.org/10.1121/1.400247>
- Fiobelkorn, I. C., Saalman, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, *23*, 2553–2558. doi:<https://doi.org/10.1016/j.cub.2013.10.063>
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*, 511–517. doi:<https://doi.org/10.1038/nn.3063>
- Hafter, E. R., & Saberi, K. (2001). A level of stimulus representation model for auditory detection and attention. *Journal of the*

- Acoustical Society of America*, 110, 1489–1497. <https://doi.org/10.1121/1.1394220>
- Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception: Entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science*, 26, 1006–1013. doi:<https://doi.org/10.1177/0956797615576533>
- Ho, H. T., Leung, J., Burr, D. C., Alais, D., & Morrone, M. C. (2017). Auditory sensitivity and decision criteria oscillate at different frequencies separately for the two ears. *Current Biology*, 27, 3643–3649. doi:<https://doi.org/10.1016/j.cub.2017.10.017>
- Holdsworth, J., Nimmo-Smith, I., Patterson, R., & Rice, P. (1988). Implementing a GammaTone filter bank [SVOS Final Report: Annex C. Part A, the Auditory Filter Bank]. Cambridge, England: Medical Research Council Applied Psychology Unit.
- Hopkins, K., & Moore, B. C. J. (2009). The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *Journal of the Acoustical Society of America*, 125, 442–446. doi:<https://doi.org/10.1121/1.3037233>
- Hsieh, I., & Saberi, K. (2010). Detection of sinusoidal amplitude modulation in logarithmic frequency sweeps across wide regions of the spectrum. *Hearing Research*, 262, 9–18. doi:<https://doi.org/10.1016/j.heares.2010.02.002>
- Hsieh, I., Petrosyan, A., Goncalves, O., Hickok, G., & Saberi, K. (2011). Observer weighting of interaural cues in positive and negative envelope slopes of amplitude modulated waveforms. *Hearing Research*, 277, 143–151. doi:<https://doi.org/10.1016/j.heares.2011.01.008>
- Hsieh, I., Fillmore, P., Rong, F., Hickok, G., & Saberi, K. (2012). FM-selective networks in human auditory cortex revealed using fMRI and multivariate pattern classification. *Journal of Cognitive Neuroscience*, 24, 1896–1907. doi:https://doi.org/10.1162/jocn_a_00254
- Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Review*, 84, 541–577. doi:<https://doi.org/10.1152/physrev.00029.2003>
- Klump, G. M., & Langemann, U. (1992). The detection of frequency and amplitude-modulation in the European starling (*Sturnus Vulgaris*)—Psychoacoustics and Neurophysiology. In Cazals, Y., Horner, K., and Demany, L. (Eds.). Book series: Advances in the biosciences (pp. 353–359, Vol. 83). Oxford: Pergamon Press Ltd.
- Lakatos, P., Musacchia, G., O’Connell, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, 77, 750–761. doi:<https://doi.org/10.1016/j.neuron.2012.11.034>
- Lamminmaki, S., Parkkonen, L., & Hari, R. (2014). Human neuromagnetic steady-state responses to amplitude-modulated tones, speech, and music. *Ear and Hearing*, 35, 461–467. doi:<https://doi.org/10.1097/AUD.0000000000000033>
- Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, 22, 1000–1004. doi:<https://doi.org/10.1016/j.cub.2012.03.054>
- Langner, G. (1992). Periodicity coding in the auditory system. *Hearing Research*, 60, 115–142. doi:[https://doi.org/10.1016/0378-5955\(92\)90015-F](https://doi.org/10.1016/0378-5955(92)90015-F)
- Langner, G., Dinse, H. R., & Godde, B. (2009). A map of periodicity orthogonal to frequency representation in the cat auditory cortex. *Frontiers in Integrative Neuroscience*, 16. doi:<https://doi.org/10.3389/fneuro.07.027.2009>
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54, 1001–1010. doi:<https://doi.org/10.1016/j.neuron.2007.06.004>
- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user’s guide* (2nd). New York, NY: Psychology Press. doi:<https://doi.org/10.4324/9781410611147>
- Meddis, R., Hewitt, M. J., & Shackleton, T. M. (1990). Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse. *Journal of the Acoustical Society of America*, 87, 1813–1816. doi:<https://doi.org/10.1121/1.399379>
- Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Molholm, S. (2013). Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory interaction. *NeuroImage*, 79, 19–29. doi:<https://doi.org/10.1016/j.neuroimage.2013.04.060>
- Morimoto, T., Irino, T., Harada, K., Nakaichi, T., Okamoto, Y., Kanno, A., ... Ogawa, K. (2019). Two-point method for measuring the temporal modulation transfer function. *Ear and Hearing*, 40, 55–62. doi:<https://doi.org/10.1097/AUD.0000000000000590>
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, 3, 320. doi:<https://doi.org/10.3389/fpsyg.2012.00320>
- Peters, R. W., Moore, B. C. J., & Baer, T. (1998). Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *Journal of the Acoustical Society of America*, 103, 577–587. doi:<https://doi.org/10.1121/1.421128>
- Re, D., Inbar, M., Richter, C. G., & Landau, A. N. (2019). Feature-based attention samples stimuli rhythmically. *Current Biology*, 29, 693–699. doi:<https://doi.org/10.1016/j.cub.2019.01.010>
- Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, 22, 807–813. doi:<https://doi.org/10.1016/j.cub.2012.03.025>
- Saberi, K., & Hafter, E. R. (1995). A common neural code for frequency- and amplitude-modulated sounds. *Nature*, 374, 537–539. doi:<https://doi.org/10.1038/374537a0>
- Schlauch, R. S., & Hafter, E. R. (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *Journal of the Acoustical Society of America*, 90, 1332–1339. <https://doi.org/10.1121/1.401925>
- Schreiner, C. E., & Malone, B. J. (2015). Representation of loudness in the auditory cortex. *Handbook of Clinical Neurology*, 129, 73–84. doi:<https://doi.org/10.1016/B978-0-444-62630-1.00004-4>
- Scott, D. M., & Humes, L. E. (1990). Modulation transfer functions: a comparison of the results of three methods. *Journal of Speech and Hearing Research*, 33, 390–397. doi:<https://doi.org/10.1044/jshr.3302.390>
- Sek, A., & Moore, B. C. J. (2003). Testing the concept of a modulation filter bank: The audibility of component modulation and detection of phase change in three-component modulators. *Journal of the Acoustical Society of America*, 113, 2801–2811. doi:<https://doi.org/10.1121/1.1564020>
- Simon, D. M., & Wallace, M. T. (2017). Rhythmic modulation of entrained auditory oscillations by visual inputs. *Brain Topography*, 30, 565–578. doi:<https://doi.org/10.1007/s10548-017-0560-4>
- Slaney, M. (1998). Auditory toolbox: A MATLAB toolbox for auditory modeling work (Technical Report 1998-010). Palo Alto, CA: Interval Research Corporation.
- Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *Journal of Neuroscience*, 34, 3536–3544. doi:<https://doi.org/10.1523/JNEUROSCI.4385-13.2014>
- ten Cat, C., & Spierings, M. (2019). Rules, rhythm and grouping: Auditory pattern perception by birds. *Animal Behavior*, 151, 249–257. doi:<https://doi.org/10.1016/j.anbehav.2018.11.010>

- Thorne, J. D., De Vos, M., Campos Viola, F., & Debener, S. (2011). Cross-modal phase reset predicts auditory task performance in humans. *Journal of Neuroscience*, *31*, 3853–3861. doi:<https://doi.org/10.1523/JNEUROSCI.6176-10.2011>
- VanRullen, R., & Macdonald, J. S. P. (2012). Perceptual echoes at 10 Hz in the human brain. *Current Biology*, *22*, 995–999. doi:<https://doi.org/10.1016/j.cub.2012.03.050>
- Wright, B. A., & Fitzgerald, M. B. (2017). Detection of tones of unexpected frequency in amplitude-modulated noise. *Journal of the Acoustical Society of America*, *142*, 2043–2046. doi:<https://doi.org/10.1121/1.5007718>
- Xiang, J. J., Peoppel, D., & Simon, J. Z. (2013). Physiological evidence for auditory modulation filterbanks: Cortical responses to concurrent modulations. *Journal of the Acoustical Society of America*, *133*, EL7–EL12. doi:<https://doi.org/10.1121/1.4769400>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.