

Temporal predictability enhances auditory detection

Emma L. A. Lawrance, Nicol S. Harper, James E. Cooke,
and Jan W. H. Schnupp

*Department of Physiology, Anatomy and Genetics, University of Oxford,
Sherrington Building, Parks Road, Oxford OX1 3PT, United Kingdom
emma.lawrance@univ.ox.ac.uk, nicol.harper@dpag.ox.ac.uk, james.cooke@dpag.ox.ac.uk,
jan.schnupp@dpag.ox.ac.uk*

Abstract: Periodic stimuli are common in natural environments and are ecologically relevant, for example, footsteps and vocalizations. This study reports a detectability enhancement for temporally cued, periodic sequences. Target noise bursts (embedded in background noise) arriving at the time points which followed on from an introductory, periodic “cue” sequence were more easily detected (by ~ 1.5 dB SNR) than identical noise bursts which randomly deviated from the cued temporal pattern. Temporal predictability and corresponding neuronal “entrainment” have been widely theorized to underlie important processes in auditory scene analysis and to confer perceptual advantage. This is the first study in the auditory domain to clearly demonstrate a perceptual enhancement of temporally predictable, near-threshold stimuli.

© 2014 Acoustical Society of America

PACS numbers: 43.66.Lj, 43.66.Mk, 43.66.Fe [QJF]

Date Received: December 13, 2013 Date Accepted: April 24, 2014

1. Introduction

Many natural auditory stimuli contain temporal regularities,¹ for instance, footsteps of potential prey or predators, or the structure of the speech envelope. These regularities make the on-going stimuli partly predictable, which is a feature that the brain may be able to exploit.² Predictive coding is a theoretical framework whereby the brain exploits regularities in sensory input to optimize processing.³ Evidence for this to date comes largely from the visual domain.⁴ In comparison, the use of predictive coding to process temporal regularities in audition is a relatively understudied phenomenon. Here we investigate the possibility that the detectability of near threshold acoustic events (noise bursts) might be enhanced if these events form a continuation of a supra-threshold periodic series.

One neural process thought to underlie anticipatory processes is “entrainment,” or “phase-locking” of oscillations in activity (membrane potential or firing rates) of a neuronal ensemble to the structure of a periodic, statistically predictable attended stimulus.⁵ Specifically, low-frequency oscillations in cortical neural activity (including auditory cortex^{6,7}), which reflect cyclical changes in population excitability,⁸ can align with the timing of event onsets.⁴ Low-frequency neuronal oscillations can become entrained to the envelope of natural sounds⁹ and speech.¹⁰ Hence entrainment concentrates high-excitability phases of activity at times of expected important auditory events, increasing their likelihood of eliciting a neural response.^{7,11}

This physiological entrainment of fluctuations in neural excitability should have perceptual and behavioral consequences if it is of functional importance. Indeed, temporal regularity in visual stimuli has been shown to improve target discrimination.² In the auditory domain, stimulus regularity and/or entrainment have been theorized to influence sound source segregation (“the cocktail party problem”)¹ and attention (dynamic attending theory¹²⁻¹⁴). Various experiments show enhanced detection of

pitch,¹⁵ spectral alterations in a sequence of tones,¹⁶ or altered interval durations^{17,18} for targets arriving in line with expectations conveyed through a periodic sequence. Temporal expectations set up through auditory regularities also decrease reaction times to targets¹⁹ and alter attention to deviant stimuli.²⁰ Indeed, behavioral measures of the influence of temporal expectation have been linked to aggregate electrophysiological measurements (electroencephalography, EEG) of entrainment: Improved reaction speed is observed when target stimuli coincide with the peak of the entrained EEG pattern.²¹ However, the effect of auditory regularity on near-threshold target detection is little examined. Ng *et al.*²² reported a neuronal phase modulation of detection that was stronger for “miss” than for “hit” trials, suggesting an inhibiting, but not “ensuring” role of entrainment on detection. It has been argued that their “entraining” stimulus did not include sufficient temporal structure to promote entrainment.¹¹ Additionally, a gap detection study suggested a role for neuronal oscillatory phase in modulating gap detection.²³ In apparent contrast to work from the visual domain, Zoefel and Heil²⁴ did not find the detectability of near-threshold tones presented at 0.5 Hz to depend on the phase of delta-oscillations, although the amplitude of the 0.5 Hz entrained oscillation was higher for hits than for misses. As yet, no study has explored the influence of continuing periodicity on auditory detection. We investigate this, against a backdrop of strong ethological relevance and theorized likelihood that regularity modifies auditory processing.

This psychophysics study examines whether there are decreased detection thresholds for temporally predictable stimuli following an explicit temporally periodic “entraining” introductory sequence. The use of target stimuli near threshold increases the perceptual demand, allowing us to explore perceptual modulations of “entrainment,” reminiscent of Cravo *et al.*⁴ in the visual domain. We compared the detection thresholds for five “on-beat” periodic target noise-bursts, to those of five aperiodic but otherwise identical targets. If the temporal predictability (induced by the introductory sequence) of periodic targets enhances their processing, then periodic targets should be detectable at lower signal-to-noise ratios (SNRs). This was indeed observed, with significantly enhanced detection of the periodic stimuli relative to the aperiodic stimuli.

2. Methods

The methodology was approved by the Ethical Review Committee of the Experimental Psychology Department of the University of Oxford, and it conforms to the ethical standards of the 1964 Convention of Helsinki. We used an adaptive method to determine detection thresholds for periodic and aperiodic broadband noise-bursts in a background of continuous broadband noise. Experiments were conducted in a brightly illuminated sound-attenuating chamber. MATLAB was used for stimulus generation, response collection, and analysis. Stimuli were played through a TDT RM1 mobile processor (Tucker Davis Technologies, Alachue, FL, USA), and presented diotically over Sennheiser (Wedemark, Germany) HD 650 headphones.

All stimuli were digitally generated broadband Gaussian noise with a sampling rate of 48 828 Hz. Background noise was presented at a comfortable listening level (55 dB sound pressure level) for the duration of the experiment. All stimuli were constructed as sequences of brief (25 ms) noise bursts atop the on-going noise background. These were implemented by transiently raising the noise intensity by a given value L in dB (hence an L of 0 dB signifies the lack of a noise burst). Signal intensities are reported in dB signal-to-noise ratio (SNR). The relationship between the noise burst level above background, L (dB), and the signal-to-noise ratio, R (in dB) is given by the equation

$$R = 10 \log_{10}(10^{(L/10)} - 1). \quad (1)$$

All stimuli started with an introductory section of seven noise-bursts presented at a regular 4 Hz rate, but with rapidly decreasing intensities (8.6389, 6.9732, 5.4554,

4.1244, 3.0103, 2.1244, 1.4554 dB SNR) to generate the percept of a periodic signal that rapidly fades into the background noise. The introductory section of seven noise bursts was followed by a target section which either contained a further five target bursts (25 ms, all with identical SNR) or it did not. Stimuli were presented in pairs, one with target noise bursts and one without, in random order (with equal likelihood), and participants were asked to indicate which of the two stimulus intervals contained target bursts (a “Two Interval Forced Choice” or 2IFC procedure).²⁵ To determine detection thresholds, the sound levels of the target bursts were adjusted from trial to trial using an adaptive procedure described further below. During any one 2IFC trial, the five target bursts could be either periodic or aperiodic (Fig. 1). For periodic stimuli, the five target bursts followed the 4 Hz rate established by the seven introductory bursts, so their timing was predictable. In aperiodic stimuli, in contrast, the intervals between target bursts were independently and randomly chosen from a uniform random distribution over 150–350 ms, so the 4 Hz time frame established by the introductory noise bursts provided almost no cues as to when the target bursts might occur. Trials containing either periodic or aperiodic stimuli were randomly interleaved (with equal likelihood), so that detection thresholds could be determined for each stimulus type. No performance feedback was provided.

The target burst intensity started at 3.35 dB SNR for both conditions (very easily audible, with consequent perfect performance) and was reduced in 2 dB SNR steps for each correct response until the first participant error. Thereafter, target burst SNR was adjusted depending on performance in the previous two trials according to the following schedule: correct-correct: -0.5 dB SNR, incorrect-correct: no change, correct-incorrect: $+0.5$ dB SNR, incorrect-incorrect: $+0.5$ dB SNR. The task finished after 15 reversal points (local maxima/minima) on both adaptive tracks, with both conditions running until completion point for the lagging track.

Traditionally, adaptive methods are analyzed using thresholds calculated from the average of reversal points (burst SNR at local maxima/minima).²⁵ Such methods use only a small proportion of available data. To exploit the full data set, we used a

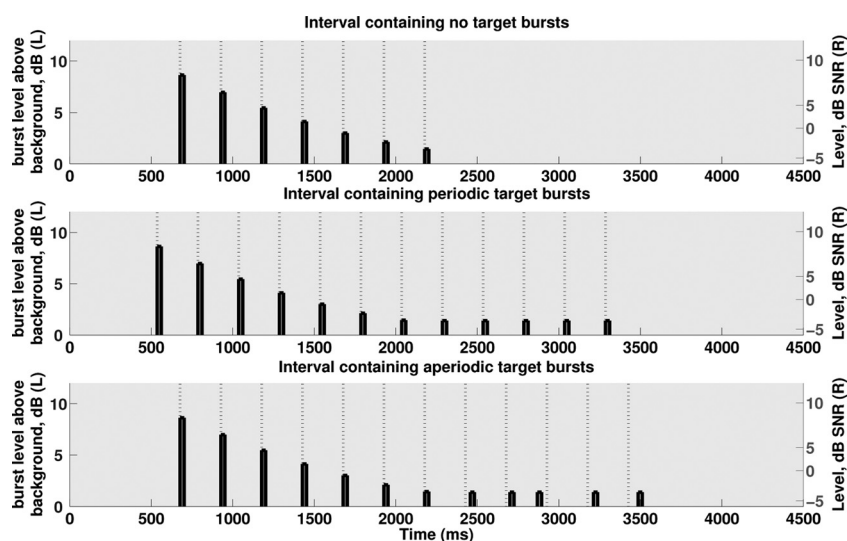


Fig. 1. Experimental protocol. Each trial consisted of two intervals (4.5 s each), with continuous broadband background noise throughout. One interval contained only seven introductory start bursts, while the other interval contained the seven introductory bursts followed by five target bursts (level adapted to participant performance at detecting target bursts). For periodic trials, the target bursts continued the introductory periodic sequence (on-beat), while for aperiodic trials the target bursts were “jittered” from the “on-beat” position. Timings shown are taken from an exemplar experimental protocol.

binomial probit regression analysis to fit psychometric curves to the responses to every trial. These curve fits were further constrained to chance (50%) performance at target burst levels of $L=0$ dB, and to perfect (100%) performance at $L \geq 3$ dB). While it has been shown that fitting probit models to data gathered with an adaptive method introduces a positive bias in the estimated slope parameters, constraining the model to $p=0.5$ for $L=0$ dB as we have done, has been demonstrated to reduce this bias to very small values (a few percent of the slope for data sets containing 100 points or more).²⁶ The probit analysis models the relationship between the probability of a correct response and the stimulus parameters on each trial according to

$$P(\text{correct response}) = \Phi(b_1 L + b_2 TL), \quad (2)$$

where Φ is the cumulative standard normal distribution (with $\mu=0$ and $\sigma=1$), L is the target noise-burst level above background in dB, T is the trial type (0 for periodic trials, 1 for aperiodic trials), and b_1 and b_2 are the adjustable weight parameters of the model, which were estimated from the data using maximum likelihood techniques (MATLAB function *glmfit*). When the trial type is periodic, $T=0$, and hence the probability of a correct response is explained by level alone: $P(\text{correct response}) = \Phi(b_1 L)$. When $T=1$, the effect of level on the probability of a correct response is altered by b_2 : $P(\text{correct response}) = \Phi(b_1 + b_2)L$. Hence if the aperiodic condition confers an advantage, b_2 will be positive. Conversely, if our hypothesis is correct and periodic targets are more easily detected, then b_2 will be negative, as irregularity decreases the probability of a correct response at a particular target level L .

3. Results

The 26 participants (aged 21–37, mean 24.9 yr, 16 female) were all of normal hearing. As explained above, the responses of the subjects were fitted with psychometric curves of the form described in Eq. (2), in which the b_2 parameter quantifies the “advantage” of aperiodic relative to regular, periodic target noise bursts. For 21/26 participants, b_2 was negative, indicating improved detection of periodic over aperiodic targets. For the population, b_2 is robustly significantly different from zero ($p=0.0025$, sign test), in support of our hypothesis. Unsurprisingly, there is also a significant effect of target level on detection: b_1 is on average highly significantly positive across the population: ($p = 2.98 \times 10^{-8}$, sign test), indicating that correct response probability increases with level. From the probit model [Eq. (2)] the slope of performance increase with level is given by b_1 in the periodic condition (mean: 0.85) and $b_1 + b_2$ in the aperiodic condition (mean: 0.65). Equation (3) quantifies the percentage increase (S) in the slope of the psychometric function when going from an aperiodic to periodic condition,

$$S = 100 \left(\frac{b_1}{b_1 + b_2} - 1 \right). \quad (3)$$

The mean increase in slope for the periodic over the aperiodic condition is 46% across all participants, or 63% across the 21 participants with enhanced performance in periodic trials. The detectability enhancement for periodic trials is also evident in Fig. 2(A), which shows the number of correct and incorrect trials completed across all participants for both trial types at the target levels roved during the adaptive paradigm. There is a shift in the peak of the distribution of correct and incorrect trials to lower signal-to-noise levels for the periodic trial type. As target level on each trial was dependent on participant performance (adaptive paradigm), this shift indicates higher performance in periodic trials relative to aperiodic trials. A comparable shift is evident in Fig. 2(B), which plots the ratio of correct trials to total trials, calculated from Fig. 2(A).

Detection thresholds for both trial types [level for $P(\text{correct response})=0.707$]²⁵ were determined using the probit fits to the psychometric functions fitted from the regression model and are shown in Fig. 2(C). A lower threshold for periodic over

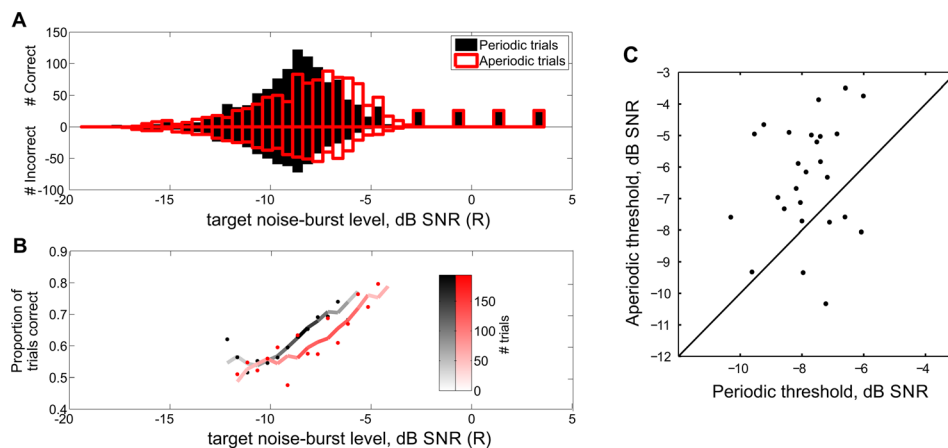


Fig. 2. (Color online) (A) Number of correct and incorrect trials for aperiodic (red) and periodic (black) trials, collated for all participants, at the target levels (dB SNR) roved during the adaptive paradigm. The target level for each trial was dependent on participant performance. A modest shift to lower signal-to-noise ratios is evident for periodic trials completed, indicative of enhanced performance relative to aperiodic trials. (B) Proportion of trials with correct target detection, for each target level bracket (three point average), with raw data also shown. The color saturation indicates the number of trials completed in each target level bracket for target levels with more than 45 trials completed across all participants. (C) Aperiodic threshold versus periodic threshold, compared to the line of unity. In 21 of 26 participants, periodic thresholds are lower than aperiodic thresholds, which indicates an enhanced detection of periodic target sounds.

aperiodic targets indicates increased detectability for the periodic targets, as observed for 21/26 participants [Fig. 2(C)]. The difference in thresholds (aperiodic-periodic) provides a measure of the enhancement conferred to the periodic stimuli. The mean difference across participants is 1.5 dB SNR (median 1.8 dB SNR), equivalent to detection of the periodic stimulus at $\sim 75\%$ of the power of the aperiodic stimulus. Hence, the temporal cueing provided by the established periodic sequence can lower detection thresholds for stimuli by approximately 1.5 dB SNR on average, although the effect size varies considerably across participants, with a lowering of thresholds by as much as 3–5 dB SNR for several participants, and with five participants showing the reverse, or no improvement

The aperiodic threshold may be thought of as a “baseline” measure of a participant’s ability to detect targets in a purely “feed forward” manner, without any advantage conferred by temporally predictable stimuli. The effect size (difference in thresholds) correlates strongly and significantly ($p < 0.0001$) with aperiodic threshold, suggesting that periodicity confers the largest absolute advantage to participants who struggle most to detect target sounds generally.

4. Discussion

Sensory processing advantages have been suggested to accrue from predictive coding for expected future events. Our study supports the hypothesis that sounds are more easily detected when they are the continuation of a temporally predictable sequence.

The variation in effect size across participants may be due to several factors, including differing entrainment susceptibility or alertness levels during the study. While there are no auditory paradigms sufficiently similar to enable an effect size comparison, an analogous recent visual experiment reports an increase of up to 55% (stimulus “intensity”) in participants’ sensitivity for unmasking visual stimuli arriving in phase with the entrained neural oscillation,²⁷ of similar order to our finding of a 25% reduction in signal power required for detection of periodic relative to aperiodic targets.

The observed perceptual advantage for periodic targets is possibly due to heightened processing of sounds arriving at high-excitability phases of a neuronal

oscillation entrained to the introductory burst sequence. The EEG can entrain to low frequency modulations of continuous noise²³ or sequences of tones.²⁰ Repetitive sounds, as used in this study, are common in behaviorally relevant scenarios, e.g., footsteps, and it has been proposed that an entrainment of neuronal oscillations can play a role in important perceptual functions such as selective attention,⁸ speech processing,²⁸ and musical processing.²⁹ The modulation rate of our stimuli (4 Hz) matches a number of modulation patterns in speech, for example, the temporal envelope of speech contains periodic structure imparted by prosody and phrasing (<4 Hz) and syllable structure (4–8 Hz, which overlaps with the theta frequencies in EEG).³⁰

Four recent studies investigated the influence of neuronal oscillatory phase on detection of auditory stimuli, but they produced mixed results.^{22–24,31} There are important distinctions between the (sometimes controversial) protocols used across these studies and our study, which may explain the heterogeneity of results.¹¹ Neuling *et al.*³¹ report that phase of alpha-oscillations entrained using oscillating transcranial direct current stimulation (o-TDCS, 10 Hz) modulates auditory target detection in background noise, with 0.6 dB normalized SNR difference between “best” and “worst” oscillatory phase. Yet there may be unknown differences between o-TDCS induced entrainment and that driven using background sound, and results could be confounded by o-TDCS periodically activating middle ear muscles.²⁴ Additionally, Henry and Obleser²³ report that gap detection varies with the phase of neural delta-oscillations entrained by a 3 Hz frequency modulated stimulus.

In contrast, Ng *et al.*²² report that the phase of “entrained” theta oscillations modulates “miss rates” for auditory target detection in natural backgrounds more strongly than “hit rates,” suggesting entrained phase can “preclude” but not “ensure” (“boost”) detection of regular auditory stimuli. Their protocol has been criticized for lacking explicit regular entrainment-promoting structure in the natural sound background.¹¹ Further Zoefel and Heil²⁴ raise the possibility that the positive results of Ng *et al.*²² and Henry and Obleser²³ may result from methodological errors in measuring entrained phase. No other study has reliably demonstrated a behavioral benefit to detectability conferred by predictable temporal regularity of an auditory stimulus.

5. Conclusions

This study demonstrates a detectability enhancement for temporally predictable near-threshold auditory stimuli. Periodic broadband noise-bursts (embedded in background broadband noise) that followed the introductory repetition rate were significantly easier to correctly detect at lower signal-to-noise levels than aperiodic bursts. The effect was modest, with a 1.5 dB SNR detection advantage conferred to the periodic stimuli relative to the aperiodic. This phenomenon is ecologically adaptive, boosting detection of relevant periodic auditory stimuli such as footsteps in background noise. Additionally, our result offers a psychophysical correlate and perceptual function for the phenomenon of neural entrainment.

Acknowledgments

This work was supported by a Clarendon Fund scholarship to E.L.A.L., and a Sir Henry Wellcome Postdoctoral Fellowship [WT076508AIA], and funds from the Department of Physiology, Anatomy and Genetics, University of Oxford to N.S.H. Thanks to Geoff Nicholls and Ben Willmore for assistance with analysis and experimental setup.

References and links

- ¹A. Bendixen, E. Schroger, and I. Winkler, “I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system,” *J. Neurosci.* **29**, 8447–8451 (2009).
- ²G. Rohenkohl, A. M. Cravo, V. Wyart, and A. C. Nobre, “Temporal expectation improves the quality of sensory information,” *J. Neurosci.* **32**, 8424–8428 (2012).
- ³R. P. Rao and D. H. Ballard, “Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects,” *Nature Neurosci.* **2**, 79–87 (1999).

- ⁴A. M. Cravo, G. Rohenkohl, V. Wyart, and A. C. Nobre, “Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex,” *J. Neurosci.* **33**, 4002–4010 (2013).
- ⁵V. J. Walter and W. G. Walter, “The central effects of periodic sensory stimulation,” *Electroencephalography Clin. Neurophysiol.* **1**, 57–86 (1949).
- ⁶C. E. Schroeder and P. Lakatos, “Low-frequency neuronal oscillations as instruments of sensory selection,” *Trends Neurosci.* **32**, 9–18 (2009).
- ⁷P. Lakatos, “An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex,” *J. Neurophysiol.* **94**, 1904–1911 (2005).
- ⁸P. Lakatos, G. Karmos, A. D. Mehta, I. Ulbert, and C. E. Schroeder, “Entrainment of neuronal oscillations as a mechanism of attentional selection,” *Science* **320**, 110–113 (2008).
- ⁹B. S. W. Ng, N. K. Logothetis, and C. Kayser, “EEG phase patterns reflect the selectivity of neural firing,” *Cerebral Cortex* **23**, 389–398 (2013).
- ¹⁰A.-L. Giraud and D. Poeppel, “Cortical oscillations and speech processing: Emerging computational principles and operations,” *Nat. Neurosci.* **15**, 511–517 (2012).
- ¹¹M. J. Henry and B. A. Herrmann, “Precluding role of low-frequency oscillations for auditory perception in a continuous processing mode,” *J. Neurosci.* **32**, 17525–17527 (2012).
- ¹²M. R. Jones and M. Boltz, “Dynamic attending and responses to time,” *Psychol. Rev.* **96**, 459–491 (1989).
- ¹³M. R. Jones, “Time, our lost dimension: Toward a new theory of perception, attention, and memory,” *Psychol. Rev.* **83**, 323–355 (1976).
- ¹⁴C. Drake, M. R. Jones, and C. Baruch, “The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending,” *Cognition* **77**, 251–288 (2000).
- ¹⁵M. R. Jones, H. Moynihan, N. MacKenzie, and J. Puente, “Temporal aspects of stimulus-driven attending in dynamic arrays,” *Psychol. Sci.* **13**, 313–319 (2002).
- ¹⁶J. M. Klein and M. R. Jones, “Effects of attentional set and rhythmic complexity on attending,” *Perception Psychophys.* **58**, 34–46 (1996).
- ¹⁷R. Barnes and M. R. Jones, “Expectancy, attention, and time,” *Cognit. Psychol.* **41**, 254–311 (2000).
- ¹⁸J. D. McAuley and M. R. Jones, “Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing,” *J. Exp. Psychol. Human Percept. Perf.* **29**, 1102–1125 (2003).
- ¹⁹J. Rimmele, H. Jolsvai, and E. Sussman, “Auditory target detection is affected by implicit temporal and spatial expectations,” *J. Cogn. Neurosci.* **23**, 1136–1147 (2011).
- ²⁰M. Schmidt-Kassow, R. I. Schubotz, and S. A. Kotz, “Attention and entrainment: P3b varies as a function of temporal predictability,” *Neuroreport* **20**, 31–36 (2009).
- ²¹G. Stefanics, B. Hangya, I. Hernadi, I. Winkler, P. Lakatos, and I. Ulbert, “Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed,” *J. Neurosci.* **30**, 13578–13585 (2010).
- ²²B. S. W. Ng, T. Schroeder, and C. A. Kayser, “Precluding but not ensuring role of entrained low-frequency oscillations for auditory perception,” *J. Neurosci.* **32**, 12268–12276 (2012).
- ²³M. J. Henry and J. Obleser, “Frequency modulation entrains slow neural oscillations and optimizes human listening behavior,” *Proc. Natl. Acad. Sci.* **109**, 20095–20100 (2012).
- ²⁴B. Zoefel and P. Heil, “Detection of near-threshold sounds is independent of EEG phase in common frequency bands,” *Front. Psychol.* **4**(262) (2013).
- ²⁵W. M. Hartmann, *Signals, Sound and Sensation* (Springer, New York, 1998).
- ²⁶C. Kaernbach, “Slope bias of psychometric functions derived from adaptive data,” *Percept. Psychophys.* **63**, 1389–1398 (2001).
- ²⁷K. E. Mathewson, M. Fabiani, G. Gratton, D. M. Beck, and A. Lleras, “Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pretarget entrainment,” *Cognition* **115**, 186–191 (2010).
- ²⁸E. M. Zion Golumbic, N. Ding, S. Bickel, P. Lakatos, C. A. Schevon, G. M. McKhann, R. R. Goodman, R. Emerson, A. D. Mehta, J. Z. Simon, D. Poeppel, and C. E. Schroeder, “Mechanisms underlying selective neuronal tracking of attended speech at a ‘cocktail party,’” *Neuron* **77**, 980–991 (2013).
- ²⁹S. Nozaradan, I. Peretz, and A. Mouraux, “Selective neuronal entrainment to the beat and meter embedded in a musical rhythm,” *J. Neurosci.* **32**, 17572–17581 (2012).
- ³⁰D. Poeppel, “The analysis of speech in different temporal integration windows: Cerebral lateralization as ‘asymmetric sampling in time,’” *Speech Commun.* **41**, 245–255 (2003).
- ³¹T. Neuling, S. Rach, S. Wagner, C. H. Wolters, and C. S. Herrmann, “Good vibrations: Oscillatory phase shapes perception,” *NeuroImage* **63**, 771–778 (2012).