


Cortical temporal integration can account for limits of temporal perception: investigations in the binaural system

Ravinderjit Singh ¹ & Hari M. Bharadwaj ^{1,2,3}✉

The auditory system has exquisite temporal coding in the periphery which is transformed into a rate-based code in central auditory structures, like auditory cortex. However, the cortex is still able to synchronize, albeit at lower modulation rates, to acoustic fluctuations. The perceptual significance of this cortical synchronization is unknown. We estimated physiological synchronization limits of cortex (in humans with electroencephalography) and brainstem neurons (in chinchillas) to dynamic binaural cues using a novel system-identification technique, along with parallel perceptual measurements. We find that cortex can synchronize to dynamic binaural cues up to approximately 10 Hz, which aligns well with our measured limits of perceiving dynamic spatial information and utilizing dynamic binaural cues for spatial unmasking, i.e. measures of binaural sluggishness. We also find that the tracking limit for frequency modulation (FM) is similar to the limit for spatial tracking, demonstrating that this sluggish tracking is a more general perceptual limit that can be accounted for by cortical temporal integration limits.

¹Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA. ²Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN, USA. ³Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA, USA.
✉email: hari.bharadwaj@pitt.edu

Neural encoding of auditory temporal information occurs at various scales across the nervous system. Temporal information is encoded with microsecond precision in the auditory periphery and is progressively transformed into a rate-based code¹. The relationship between neural temporal coding limits at various levels of the pathway and human temporal perception is poorly understood. Although temporal information in the ascending auditory pathway is progressively transformed to a rate-based code, cortical neurons can also synchronize to auditory cues, albeit at lower frequencies than subcortical structures^{1–3}. The impact of this cortical synchronization on perception is poorly understood. Studies have found cortical synchronization to be useful in explaining perception of time-reversed animal vocalizations, encoding of stimulus onset, representation of auditory objects, and the coding of amplitude modulations (AMs)^{4–9}. For instance, as the modulation rates (frequencies) of sounds are increased, the elicited percept changes from that of being able to perceptually follow discrete events (i.e., the individual peaks and troughs of the modulation), to a flutter, and then to a pitch. Indeed, correlates of these qualitative changes in the percept of AM can be found in the temporal synchronization capabilities of different cortical regions^{2,3,9}. To better understand the relationship

between neural temporal coding limits and perception, here, we studied the binaural auditory system, which shows both exquisite sensitivity to microsecond binaural temporal cues, and seemingly paradoxical “sluggishness” to dynamic variations in the same cues in certain behavioral tasks¹⁰. In particular, we studied how cortical temporal processing related to perception of dynamic binaural cues. Specifically, we characterize the cortical temporal integration window for dynamic binaural cues in humans using EEG and investigate how well that physiological window can explain human perception of dynamic binaural cues. Our inability to spatially track fast binaural modulations (BM) has been called “binaural sluggishness” in the literature, because it was thought that this sluggishness phenomenon may be unique to the binaural system. Instead, here, we explore the possibility that cortical synchronization capabilities may place common neurophysiological constraints leading to sluggish processing of a range of auditory cues, including monaural AM and frequency modulations (FM).

In this investigation, BMs are applied to two binaural cues based on the temporal fine structure in sounds, namely interaural time delay (ITD) and interaural correlation (IAC). ITD is the difference between the arrival times of corresponding components of a sound in the two ears, and is useful for sound lateralization; IAC is the correlation between sounds reaching the two ears, and can directly influence the perceived spatial extent/width of a sound. Both ITD and IAC can be useful cues for listening in noisy environments^{11–13}. How binaural cues, e.g. ITD and IAC, are processed when they are dynamic, is currently not well understood. Single-unit data from the brainstem shows that cells can encode BMs in ITD and IAC in the 100s of Hz^{14–17}. Single-unit data from primary auditory cortex (A1) shows neurons can synchronize to binaural beats, a dynamic interaural phase difference (IPD), up to a median synchronization rate of about 20 Hz^{18,19}. Behaviorally, humans can detect BMs in the 100s of Hz^{15,20,21}; however, studies probing the use of binaural cues to perform a spatial task, e.g. spatial unmasking, have found that humans only benefit from low-rate BMs, below 10 Hz^{10,22–25}. These binaural unmasking studies led to the notion of ‘binaural sluggishness’ in the literature as the binaural system seemed particularly slow in comparison to human ability to detect monaural cues. However, a broader view of the literature suggests a dichotomy where tasks that rely on the percept being spatialized in nature appear slow, while simple detection of binaural fluctuations can be more than an order of magnitude faster. Indeed, dynamic spatial percepts have been anecdotally reported to become a flutter at frequencies above approximately 7–10 Hz^{15,16,21}. Thus, BMs seem to have analogous temporal perceptual limits as AMs, in that both demonstrate a qualitative switch around 7–10 Hz, Fig. 1; with BMs, the percept switches from spatialized to a mere flutter, and with AMs, the percept switches from being able to “ride” the AM (or perceive individual peaks and troughs discretely), to also perceiving a flutter. This dichotomy (fast and slow) in the behavior is accompanied by evidence of fast temporal processing in the brainstem and slower temporal processing in the cortex. However, the slower temporal processing found in cortex thus far, particularly primary auditory cortex (median synchronization limits of 20 Hz), is not as slow as spatial perception observed in human behavior^{18,19}; however, in this comparison, it should be acknowledged that human perceptual limits are being compared to non-human physiology.

We evaluate how cortical binaural synchronization in humans explains human dynamic binaural perception. A temporal analysis window is quantified using EEG with a novel binaural systems identification technique which modulates IAC or ITD with a maximum length sequence (m-seq). Using the same system identification technique, we also estimate brainstem binaural

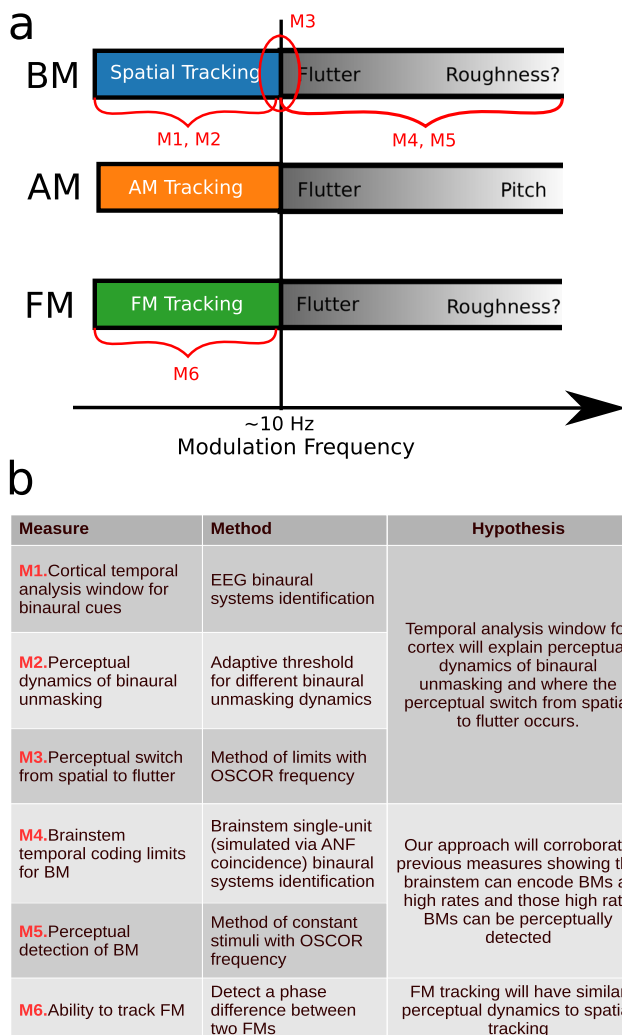


Fig. 1 Overview of experiments, methods, and hypotheses. **a** This figure depicts the similarity in temporal switches in perception that occur between binaural modulation (BM), amplitude modulation (AM), and frequency modulation (FM). In red, the measures taken in this work are depicted and laid out in the table shown in **(b)**.

synchronization limits in chinchillas to both corroborate previous brainstem measurements and to validate our novel approach against prior single-unit data obtained using conventional approaches. We conduct three behavioral experiments (1) to characterize detection limits for BMs, (2) estimate the frequency limits at which the perceptual switch from dynamic spatial to a non-spatial flutter occurs, and finally (3) to quantify the perceptual dynamics of binaural unmasking; we then compare the behavioral frequency functions to synchronization limits of cortex and brainstem. To complement our binaural measures, we also measured subjects' ability to monaurally detect a phase difference between two spectrally distant FMs, which requires being able to temporally follow the individual cycles of the FMs to test whether the ability to temporally track BMs (i.e. spatial tracking) ceases at similar modulation rates as for other auditory cues. Our results reveal a neural source with a latency of ~100 ms that synchronizes to BMs slower than 10 Hz, and can quantitatively account for the sluggish dynamics seen in spatial unmasking. We also find that BM and FM stimuli can be perceptually tracked out to similar modulation frequencies (~10 Hz). These results suggest the temporal response properties of later (~100 ms latency) areas of cortex may constrain our ability to perceptually track various dynamic auditory cues. This is also notable in that we have developed an EEG measure that measures cortical temporal coding ability which appears to have a direct relationship with temporal perception.

Results and discussion

Novel systems identification approach to measure temporal coding of binaural modulations. In the present study, we employed a novel approach to characterize neural coding of binaural modulations. Neuroscience in general, and auditory neuroscience in particular has a rich history of characterizing temporal coding. Indeed, temporal modulation transfer functions (tMTFs) have been measured from various levels of the auditory system for spectral, amplitude, and binaural modulations^{2,14,16}. The conventional approach involves playing sinusoidal modulations (i.e., a single frequency in the modulation domain), and measuring neural phase locking as a function of the modulation frequency. This approach is time consuming, as many frequencies need to be measured individually. Moreover, given the highly non-linear and adaptive nature of the central auditory neural response, the results from a discrete single-frequency sampling approach may miss interesting characteristics that are unique to the broadband nature of real-world stimuli. To mirror the complex broadband modulation profile encountered in the environment, we applied a broadband binaural modulation using a modified maximum length sequence or m-sequence (m-seq). Throughout, we refer to this modified stimulus as the “extended m-seq” (em-seq) (Fig. 2). This approach allows us to simultaneously measure the coding across all modulation frequencies of interest with one ongoing stimulus, and provides a substantial improvement in experimental efficiency compared to traditional tMTF measurements. The different experiments conducted, and hypotheses tested, are outlined in Fig. 1.

Source binaural temporal response function shows cortical tracking of binaural cues extends up to 10 Hz. Our measurement using an extended m-seq (em-seq) and subsequent PCA analysis to obtain a “source binaural temporal response function” (sBTRF) is explained in detail in the methods section. The sBTRF for IAC and ITD is shown in Fig. 3. The sBTRF reflects how the dominant underlying cortical sources respond to dynamic binaural stimuli. The PCA weights are depicted in the topomaps in Fig. 3. Considering auditory cortex has tangential dipoles that

project to the top of the scalp, the topographic distribution we find for the sBTRF for IAC and ITD is consistent with arising from auditory cortex^{26,27}. Neuroimaging data in humans and data from animals support the notion that perceived auditory space is computed in auditory regions in the temporal lobe, but beyond primary auditory cortex^{28–31}. Consistent with this, the dominant group delay (or latency) estimated from the sBTRF was ~100 ms for both IAC and ITD. Studies on the encoding of dynamic binaural cues in cortex have mainly focused on neurons in the primary auditory cortex and have reported median synchronization limits of ~20 Hz^{18,19}; the 20 Hz limit is much faster than what is observed in human behavior. The magnitude response of the sBTRF for IAC (Fig. 3) loses 6 dB or half its amplitude by 5 Hz and the mean response falls into the noise floor at ~9 Hz, while for ITD (Fig. 3), the amplitude peaks around 4.5 Hz and falls into the noise floor at ~9 Hz as well. This slower frequency limit estimated from the sBTRF further corroborates the notion that the dominant contributors may be sources that are hierarchically downstream to the primary auditory cortex.

The sBTRF shapes for IAC and ITD have an interesting relationship in that the shape of the ITD sBTRF roughly matches the derivative of the IAC sBTRF (Fig. 3). This relationship can be explained in the context of a two-channel population code model for sound location^{31–33}. With IAC, it is likely all cortical binaural cells (in both channels) respond to IAC the same way, exhibiting a smaller response at a low IAC and a larger response at high IAC. In contrast, with ITD, two-channel models considered in the literature posit that one channel would respond more favorably to sound locations on one lateral side vs the other^{31–33}. Therefore with the ITD em-seq (which bounced between two azimuths), one channel may be much more active for one azimuth and the other more active for the other azimuth leading to a derivative-like response compared to the response seen for IAC. Fig. 3 shows results from single-unit measurements in chinchillas, demonstrating that the approach utilizing an em-seq can also be readily adapted to spiking data. We estimated binaural brainstem responses from a coincidence analysis on measurements from the auditory nerve recordings in chinchillas. Previous studies have shown that brainstem response properties can be reliably predicted from nerve responses using this approach¹⁴. A drop of 6 dB in power was used as the synchronization limit to be consistent with previous literature^{14,16}. Our brainstem estimates indicate synchronization up to 100s of Hz, in line with previous measurements^{14–16}.

Spatial and frequency tracking limits align with cortical tracking of spatial cues.

Results from behavioral experiments are shown in Fig. 4. Detection thresholds for BMs were measured using a binaural oscillating-correlation (OSCOR) stimulus, where the IAC was varied sinusoidally at different BM rates. Results revealed that humans can detect BMs as fast as 100s of Hz, consistent with our ability to temporally encode fast BMs in the brainstem. Importantly, being able to *detect* BMs does not mean that the listener is making use of fluctuations in perceived lateralization for all rates. One could be simply discriminating a dynamic BM from a static sound just as one can discriminate a 500 Hz AM from a sound with no modulation by perceiving a buzz rather than 500 discrete modulations. It should be noted that experiments were primarily conducted with noise stimuli that were bandlimited between 0.2–1.5 kHz to match the range of frequencies over which humans have high sensitivity to ITDs³⁴, perhaps constrained by the limits of neural phase locking³⁵. However, we found OSCOR detection in this band-limited range was much below what we know subcortical structures are capable

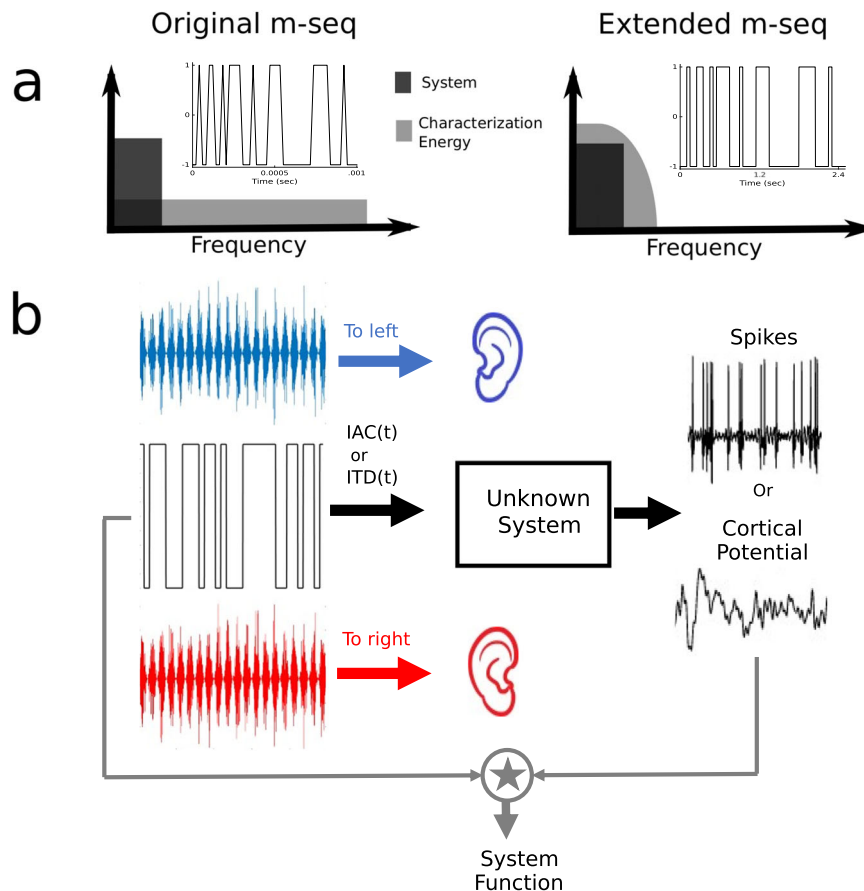


Fig. 2 Novel binaural systems identification approach utilizing m-seqs to measure binaural temporal coding. **a** depicts how the m-seq is transformed into the extended m-seq by increasing the duration of each point in the m-seq. This alters the frequency response of the m-seq to be sinc shaped instead of white, but that is useful in focusing the characterization energy in the range the system of interest is active. **b** depicts the paradigm used to obtain system responses. The extended m-seq modulated either the IAC or ITD of the noise stimulus and the neural measure, either spikes or voltage potentials (EEG), were cross correlated with the extended m-seq to obtain an estimate of the system response.

of encoding, see Fig. 4. Therefore, we repeated this experiment in one subject with white noise due to physiological data indicating cells with higher center frequencies can encode the fast OSCORs¹⁴. We found detection of the OSCOR was improved with the white noise, see Fig. 4 and Supplementary Fig. 3. Indeed, one possibility is that fine-structure-based binaural cues may be detected for higher (beyond 1.5 kHz) carriers but that these cues don't inform spatial perception. One reason for this could be due to the human head size. ITDs from higher carrier frequencies become ambiguous if the wavelength is smaller than the head width which would make fine structure cues less informative for determining spatial location at higher carrier frequencies³⁵. However computing correlation is not affected by head size, so IAC may actually be affected only by fine structure processing ability which likely extends beyond 1.5 kHz. It has been anecdotally reported that the perception of BMs switches from being spatial to a flutter between 6 and 10 Hz^{15,16,21}. We formally measured this in one participant and found the switch from spatial to flutter occurred at 9.3 Hz in that participant which is consistent with previous anecdotal reports.

In contrast to detection experiments, binaural unmasking experiments rely on subjects' ability to use brief periods of perceived spatial separation (i.e., differences in lateralization/location) between target and masking sounds to derive masking release²². Indeed, the results of our experiment measuring the dynamics of binaural unmasking reveals the sluggish nature of spatial processing; the binaural masking release (measured as dB

improvement in target detection) continues to grow even as the duration of the spatial separation is increased in the range of tenths of a second, Fig. 4. Strikingly, when we used the physiologically measured sBTRF for IAC as the temporal integration window, the predicted masking release shows excellent correspondence to the behavioral data, with the mean-squared error (MSE) of the fit being as small as 0.6 dB² (this low MSE is despite excluding the end points, i.e., 0 and the asymptotic value at 1.6 seconds which were constrained to match the measurements). Therefore, the sluggishness in binaural perception is well explained by the physiological properties of cortical processing, specifically that of a cortical source that is measurable using EEG with a group delay of ~100 ms. Data from primary auditory cortex has shown median synchronization limits to binaural cues to be 20 Hz^{18,19} which is faster than our measure here that fits the behavior, suggesting this sluggishness arises in a hierarchically later cortical source (consistent with the longer 100 ms latency) and not primary auditory cortex. Finally, to test whether a monaural percept can also exhibit sluggishness, we measured the ability to detect a phase difference between FMs at the same rate imposed on spectrally distant carriers. This task was chosen because it relies on subjects' ability to perceive discrete variations in the modulation and temporally track it (i.e., being able to "ride" the peaks and troughs) to detect the phase difference. This is analogous to being able to track the changing lateralization or spatial location with BM stimuli. We hypothesized that similar to BMs, FMs would show the same tracking

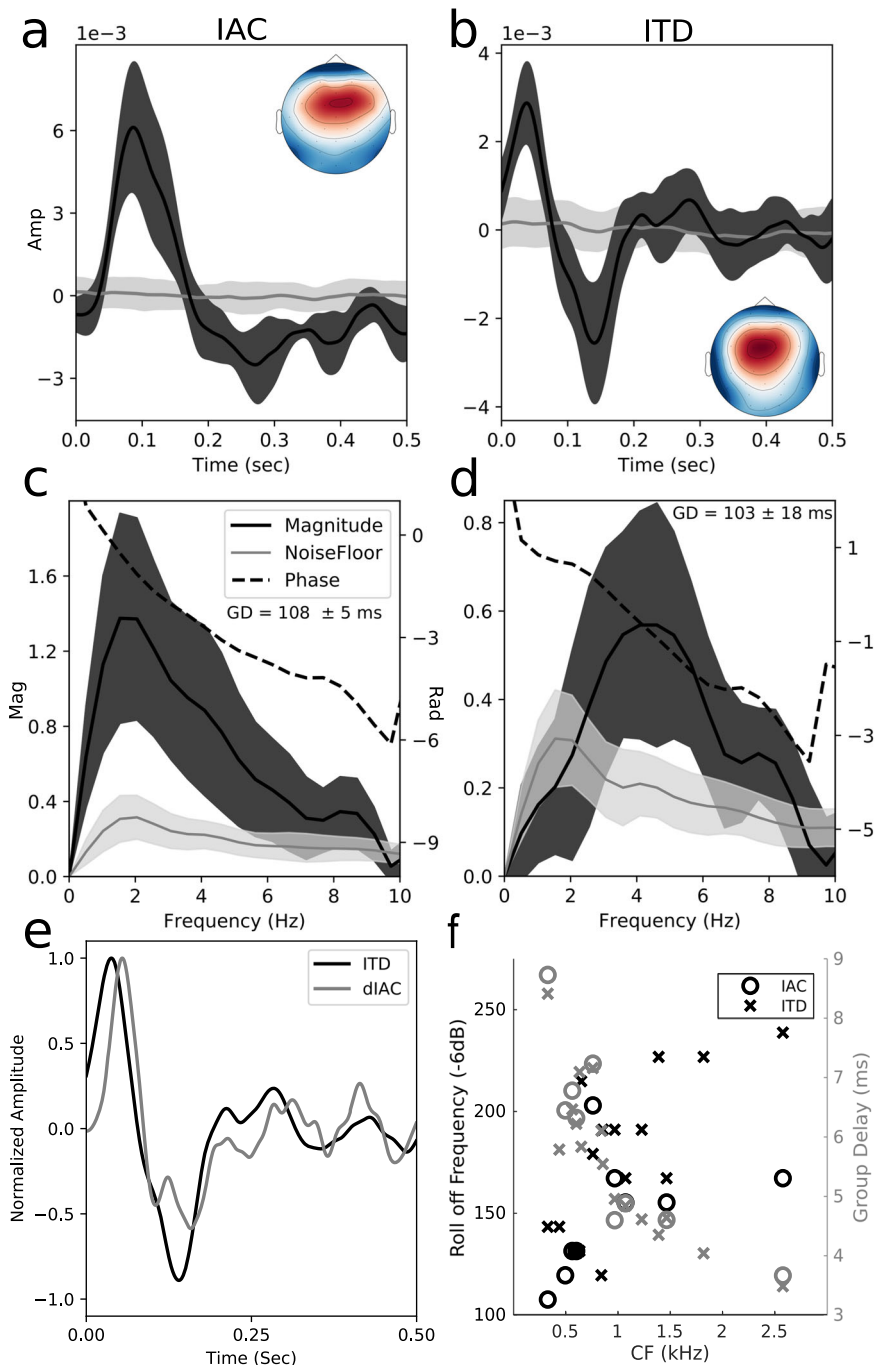


Fig. 3 Cortical tracking of IAC and ITD extends out to -10 Hz. In **a–d**, the solid line is the mean response and shading represents the 95% confidence interval calculated using the standard error computed from jackknifing ($n = 9$ independent samples). **a, b** Shows the source binaural temporal response function (sBTRF) for IAC and ITD as well as the topomap for the sBTRF. **c, d** Shows the frequency response of the sBTRF for IAC and ITD. The data was high-passed at 1 Hz. The group delay (GD) confidence interval is \pm the standard error of the estimate computed by jackknifing. **e** Shows a plot of the ITD sBTRF as well as the derivative of the IAC sBTRF with the amplitudes normalized between 1 and -1 . **f** Shows the roll off frequency and group delay for the brainstem responses simulated from ANF coincidence.

limits. Indeed, we observed that FMs can be temporally tracked up to about 10 Hz, similar to our BM results, Fig. 4.

Conclusions. In summary, the sBTRF measured for IAC and ITD provides a neural correlate for the sluggish perception of auditory space. Furthermore, the latency of the dominant source contributing to the sBTRFs suggests the perceptual limits arise from auditory cortical regions downstream of the primary auditory cortex. Previous investigations focusing on the brainstem and

midbrain processing have not produced any candidate correlates for this spatial sluggishness^{14–17,36,37}. Cortical single-unit measurement in A1 have revealed lower limits of temporal coding than the brainstem, but not low enough to account for the behavioral estimates of sluggish spatial perception in humans^{18,19}. The shape of the sBTRF measured here yielded sluggishness estimates that were a very close match to the spatial sluggishness seen in binaural unmasking. Our results also make it clear that the processing of binaural cues is not uniquely sluggish, and that monaural temporal

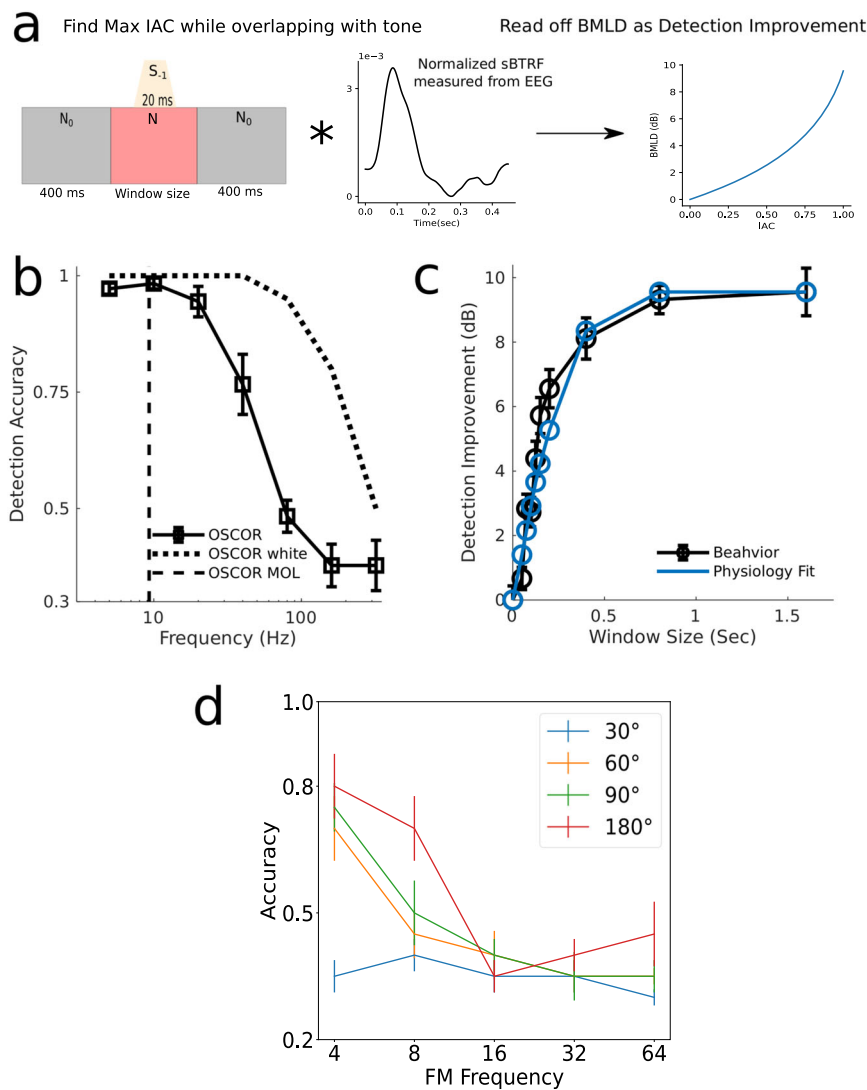


Fig. 4 Cortical temporal coding limits of binaural cues align with spatial tracking of binaural cues, spatial unmasking dynamics, and FM modulation tracking. **a** demonstrates how the physiology fit was obtained and is based on procedures from Culling and Summerfield²². The window size in the figure is the duration of time in the experiment that fully correlated noise was played with the target anticorrelated tone being in the middle of that window (see Methods section). **b** shows results from all experiments using the OSCOR stimulus. OSCOR white ($n = 1$) is the OSCOR stimulus with the noise having a white spectrum (instead of the noise being band-limited to 0.2–1.5 kHz, $n = 9$ and bars show \pm standard error across subjects) and OSCOR MOL is the result of the method of limits experiment with the OSCOR stimulus, $n = 1$. **c** shows the results for a binaural unmasking paradigm as well as how well the IAC sBTRF window fits the behavioral data ($n = 9$ for behavior, bars show \pm standard error across subjects). The window size is the same window size as in **a**. **d** shows the results from our FM phase difference detection experiment ($n = 15$, bars show \pm standard error across subjects). Box plots to visualize the full distribution of data for **b**, **c** and **d** can be found in Supplementary Fig. 6.

perception also exhibits sluggishness when the behavioral tasks are set up to probe analogous aspects of temporal tracking. This suggests that “binaural sluggishness” may be somewhat of a misnomer. Indeed, with AM, fluctuations can be followed only up to about 10 Hz if individual acoustic events are to be discretely perceived; then at higher rates, an acoustic flutter is heard (similar to the binaural flutter), and then finally a pitch⁹. Correlates of these qualitative changes in AM perception can be found in the temporal coding properties of different regions along the auditory pathway^{3,9}. Here, we demonstrated that a similar 10 Hz limit also exists in the tracking of FMs. Indeed, given that hierarchical cortical processing occurs via highly conserved columnar circuits³⁸, and that cortical neurons generally exhibit tuning to a range of different acoustic cues at the single-unit level, it is reasonable to expect that the temporal coding properties of cortical neurons would constrain the processing of many different cues similarly. In this view, the

sluggishness in tracking a moving sound (spatial sluggishness) would be no different than the sluggishness in tracking discrete amplitude or frequency fluctuations. Rather, cortical processing limits may impose a general sluggishness in the processing of a range of auditory cues. Our results also suggest that although the temporal integration window progressively expands as we ascend the auditory pathway (progressive temporal to rate transformation), the temporal synchronization properties of higher stages of cortical processing can still directly relate to certain aspects of perception. The perceptual significance of this low frequency temporal processing is understudied, and could be important for our understanding of the mechanisms supporting the perception of complex acoustic scenes. Our em-seq approach to evaluate how well cortex can track a particular auditory cue can be readily extended to other acoustic cues, such as AM, to investigate cortical temporal coding more broadly. Finally, to the extent that temporal

response properties of higher-order areas relate to perception of dynamic scenes, individual differences in physiology measured using the em-seq approach may also be useful for predicting individual listening outcomes in complex everyday environments. Future investigations evaluating how cortical temporal coding ability measured via EEG with an em-seq relates to speech-in-noise processing, aging, and other phenomenon could produce interesting lines of investigation to better understand the significance of cortical temporal processing.

Methods

We conducted six experiments, illustrated in Fig. 1, to investigate how temporal processing in cortex relates to binaural perception and to evaluate whether a common sluggishness phenomenon can be found across the processing of many auditory cues. Two experiments involved electrophysiological measurements using our novel binaural systems identification approach to quantify the temporal coding properties of brainstem and cortex for binaural modulations. Four experiments were behavioral studies in humans designed to probe the temporal limits of different binaural perceptual abilities, and to test whether physiologically measured limits can predict perception.

Human Participants. Nine participants (2 female) with an average age of 25 (18–34) were recruited from the greater Lafayette area through posted flyers and advertisements. Audiograms were measured using calibrated Sennheiser HDA 300 headphones by employing a modified Hughson-Westlake procedure. All subjects had hearing thresholds better than 25 dB HL in both ears at octave frequencies between 250 and 8 kHz. All subjects provided informed consent, and all measurements were made in accordance with protocols approved by the Internal Review Board, and the Human Research Protection Program at Purdue University.

EEG recording. Digital stimuli were designed with custom scripts in Matlab (The MathWorks Inc., Natick, MA) at a sampling rate of 48828.125 Hz, and converted to analog voltage signals using an RZ6 audio processor (Tucker-Davis Technologies, Alachua, Florida). The voltage signals were converted to sounds and delivered to the ears via ER2 insert earphones (Etymotic Research, Elk Grove Village, IL) coupled to foam ear tips. There was a random jitter between 0–200 ms added to each interstimulus interval to reduce any potential periodic noise sources that could be in phase with our stimulus. EEG measurements were made with a 32-channel system (Biosemi Active II system, Biosemi, Amsterdam, Netherlands). EEG data was sampled at 4096 Hz and filtered between 1–40 Hz, and then downsampled to 2048 Hz. EEG data were re-referenced to the average of all electrodes. Ocular artifacts were removed using the signal-space projection approach³⁹. Projections were manually applied for each participant based on the topographic pattern of the noise-space weights which were manually chosen to correspond to the expected pattern of blink and saccade artifacts. EEG data were collected as stimuli were played passively with participants watching a muted video of their choosing with subtitles. To remove movement artifacts, trials that had peak to peak deflections greater than 200 μ V were then rejected.

Auditory nerve responses to simulate binaural processing. Auditory nerve responses were collected in chinchillas. Male chinchillas ($N = 5$) weighing 400–650 grams under 2 years of age were used in accordance with protocols approved by the Purdue Animal Care and Use Committee. Binaural processing was simulated by playing the stimulus intended for each ear separately and recording from the same nerve fiber, and then performing a

coincidence analysis. Coincident spikes (within 50 μ s bins) across the “left” and “right” spike train was treated as amounting to a binaural spike. Previous work has shown that this coincidence analysis approach can reliably estimate the response properties of binaural cells in the midbrain¹⁴. All procedures laid out in this section were approved by the Purdue University Animal Care and Use Committee.

Anesthesia was induced with xylazine (1–2 mg/kg s.c.) and ketamine (60–65 mg/kg, s.c.). Anesthesia was maintained with ketamine (20–40 mg/kg, i.m.) and diazepam (1–2 mg/kg, i.m.) through intramuscular injections every 2 h as indicated by the presence of reflexes for the duration of the experiment (10–16 h). A heating blanket and rectal thermometer were used to regulate and monitor body temperature throughout the experiment. The skin and muscles overlying the skull were reflected to expose the bony ear canals and bullae. Hollow ear bars were placed close to the tympanic membrane. The AN bundle was exposed at its exit from the internal acoustic meatus via a posterior-fossa craniotomy and aspiration cerebellotomy. Acoustic stimuli were presented monaurally through an ear bar using Etymotic ER2 and calibrated using a probe microphone placed within a few mm of the tympanic membrane (Bruel and Kjaer 4182). Glass pipettes with impedance ranging between 60–90 M Ω were advanced into the auditory nerve using a hydraulic microdrive. Recordings were amplified, band-pass-filtered and stored on a PC. Spikes were identified using a time-amplitude window discriminator, and spike times were stored with 10 μ s resolution.

Single fibers were isolated by monitoring the spike recording via an oscilloscope and listening for spikes while noise pips were played. After isolating a fiber, its tuning curve was measured using an automated procedure that played a series of tone pips and measured the minimum level required to evoke 1 more spike than a subsequent 50 ms silent period⁴⁰. Units where we were unable to measure at least 20 repetitions of the left and right stimulus were excluded from analysis.

Novel binaural systems identification using maximum length sequences. Systems identification approaches are widely used in Neuroscience^{41,42}. In addition, systems identification approaches are useful for determining the frequency responses of a system, both at single-neuron and population levels^{43,44}. Binaural systems identification has been used to study the temporal coding limits of the binaural system primarily through temporal modulation transfer functions (tMTF) computed from calculating phase locking to individual frequencies of binaural modulation^{14–16,18,19}. However, the approach of constructing a tMTF from responses to individual frequencies is experimentally inefficient. Also central neural responses can be sensitive to the input statistics of a stimulus, so an approach that better mimics the broadband nature of real-world stimuli may capture interesting characteristics that a discrete single-frequency sampling approach may miss^{45,46}. Here we use a novel binaural systems identification technique where we modulate a binaural cue using a maximum length sequence (m-seq) allowing us to evaluate a continuous range of frequencies using a single stimulus.

Maximum length sequences (m-seqs) have been used to map the receptive fields of visual neurons, identify the cues used in visual tracking, obtain auditory brainstem responses, and characterize room acoustics^{47–51}. An m-seq is a signal that only takes one of two discrete values (e.g., +1 or –1). It is a pseudorandom sequence constructed through a series of feedback shift registers of n bits, giving the full sequence a length of $2^n - 1$. An m-seq has a flat spectrum, similar to white noise, but has a lower crest factor in that its values are bounded on both ends, making it attractive for systems identification. By using the

m-seq as the input to a system and then cross-correlating it with the output of the system, an estimate of the impulse response can be obtained. For more details on the m-seq, we suggest the following sources^{47–50}.

The m-seq has a nearly white spectrum; however auditory responses to binaural modulation, especially in cortex are expected to be slow. Indeed, previously measures from the primary auditory cortex of rabbits and macaques suggest that responses roll off around 20 Hz^{18,19}. Thus, if we played the m-seq as it is conventionally constructed, its energy would be spread out up to $f_s/2$, which is over 20 kHz in this case, because our sampling rate (f_s) is 48828.125 Hz. This is orders of magnitude wider than expected frequency response range of the system resulting in much of the stimulus energy for system characterization being spent without eliciting a measurable response component, leaving only a small fraction to characterize the system in the region that it is active (e.g., below 20 Hz). Therefore we modified the m-seq to an extended m-seq (em-seq) to have a sinc-function-shaped spectrum instead of a white spectrum, so that most of the characterization energy of the m-seq is in the region of interest for the system being characterized. The sinc-function shape will modulate the measured system response; however, the em-seq can be designed so that its frequency response is mostly flat in the region the underlying system is active. The procedure to construct the em-seq is explained next.

The em-seq is constructed by elongating the duration of each point in the conventional m-seq to create an extended m-seq, i.e., em-seq. In the conventional construction, each point in the m-seq is $1/f_s$ in duration where f_s is the sampling rate. By elongating the duration of each sample to T , the em-seq spectrum instead of being flat, takes a sinc-function shape, with only minimal energy above $f = 1/T$, but approximately flat (losing <4 dB) between $f = 0$, and $f = 1/2T$. For example, if characterizing a system that is expected to be active up to 40 Hz, elongating each point to a duration of 12.5 ms would be appropriate because then the resulting sinc spectrum will be fairly flat up to 40 Hz and lose energy quickly between 40–80 Hz. The frequency at which the main lobe of the em-seq spectrum loses nearly all power will be denoted as the cutoff frequency (COF; e.g., 80 Hz in the previous example). The frequency up to which the em-seq spectrum is approximately flat (i.e., loses ~4 dB in power) will be denoted as f_{ADB} , which is 40 Hz in the previously given example. Given the properties of a sinc function, $COF = 2 * f_{ADB}$. A COF or f_{ADB} is set by choosing the elongation duration applied to the conventional m-seq to obtain the em-seq. When choosing the f_{ADB} , the idea is to have most if not all of the system's energy that is being characterized to be below this number. However, if the system has energy beyond f_{ADB} , that will be captured because the em-seq still has energy above f_{ADB} . However, the system has a higher chance of being shaped by the spectrum of the em-seq itself. To best capture the true response of the system, it is important to establish noise floors that help determine where the system is characterizable and where the measurement is just noise. We describe our approach to do this with EEG data and spiking data later in this paper. Lastly, the number of bits for the em-seq is a parameter that will need to be chosen, as is the case with a conventional m-seq. The minimum requirement is that the total length of the em-seq needs to be longer than the expected impulse response.

Two different em-seqs were used for characterizing brainstem (via ANF coincidence) and cortical responses (via EEG) given that the two systems were expected to reflect very different temporal coding abilities. The brainstem em-seq was a 9 bit m-seq with $T = 2$ ms, so f_{ADB} was 250 Hz. A minority of units were measured with the elongation parameter set at 1 ms in duration, i.e., f_{ADB} : 500 Hz; the data from these units indicated that

f_{ADB} of 250 Hz was sufficient. The brainstem em-seq was 1.026 s in duration and was presented 6 times. The cortical em-seq was 8 bits with each point being elongated to 50 ms, giving a f_{ADB} of 10 Hz. The 50 ms duration for each em-seq point was chosen after piloting using f_{ADB} of 20 Hz suggested cortical phase locking limits to binaural modulations measured with EEG were below 10 Hz. Also from our results (see Fig. 3), it is clear that the em-seq we choose was adequate. At 6 Hz for the IAC response, the magnitude has dropped by 1/3 or roughly 10 dB, but the em-seq used here has dropped only by ~1.5 dB (see Supplementary Fig. 5), so the response is losing its energy much faster than the em-seq. If the response had energy beyond f_{ADB} , we would capture that because the response likely would not fall into the noise floor and continue beyond f_{ADB} as the em-seq has energy beyond f_{ADB} . However, if the measured response has substantial energy beyond the chosen f_{ADB} , that would be an indication to modify the em-seq to have a larger f_{ADB} . The total duration of the cortical em-seq was 12.75 s, and we aimed to collect 300 trials for BMs applied to ITD and IAC. The interstimulus interval was 1 s plus a random jitter between 0 and 200 ms. One participant did not complete all trials for the ITD em-seq, and one participant did not complete all trials for the IAC em-seq due to availability constraints.

Figure 2 depicts the paradigm used in this study. An em-seq modulated either the ITD or IAC of a broadband noise stimulus. The noise stimulus used for single-unit data had a bandwidth of 0.01–20 kHz and 0.2–1.5 kHz for the cortical stimulus. The bandwidth for the cortical stimulus was narrower since humans do not appear to use fine structure binaural cues for spatial perception beyond ~1.5 kHz³⁴. An amplitude modulation (AM) was imposed at the highest frequency the em-seq can characterize (so 500 Hz for the brainstem em-seq and 20 Hz for the cortical; matching the fastest rate at which the em-seq was allowed to switch states) to mask any phase discontinuities introduced by jumping between two ITD or IAC values. This does not affect our analysis because the em-seq would have lost most of its power at this frequency. The modulation by the em-seq was binaural, so the em-seq binaural modulation could only be heard with both earphones in place. Listening with just one earphone would lead to perceiving just 20-Hz amplitude modulated noise. The m-seq for IAC bounced between an IAC of 1 and -1 , and for ITD bounced between 0 and 500 μ s for EEG data collection. For single-unit data, the ITD bounced between 0 and $1 / (2 * CF)$, where CF is the characteristic frequency of the unit being measured. This was chosen based on the observation that ITD tuning curves of brainstem single neurons exhibit a trough at an ITD that is a phase of π from the CF of the unit⁵², and because we would expect greatest probability of coincidence between left and right auditory nerve responses for an ITD of 0, and minimum coincidence at a ITD corresponding to a phase difference of π at CF.

Em-seq analysis. The em-seq used for analysis will be termed the recovery em-seq (rem-seq). The rem-seq always goes between 1 and -1 , so with IAC, the em-seq played as the stimulus and the rem-seq are the same. For ITD, however, the em-seq bounces between two different ITD values, but for analysis the rem-seq goes between 1 and -1 . This choice of rem-seq avoids introducing an artificial DC value into the analysis, and an overall scaling that would not affect the frequency shape of the system response, thus making it well-suited to characterize how the system can track modulations. After a response is cross-correlated with the rem-seq, it is unitless; therefore, many of our responses are unitless. However, since we are primarily interested in the shape of the responses, having unitless responses is fine.

For the single-unit data, the spikes were treated as Dirac delta functions (i.e., single-sample impulses indicating that a spike occurred in a narrow time bin of the peristimulus time histogram) and cross-correlated with the rem-seq to extract the binaural impulse response, i.e. binaural temporal response function. Noise floor estimates were constructed by randomizing the inter-spike intervals (ISIs) and making the first spike time be drawn from a uniform distribution between 0 and the actual first spike time for that trial, and then doing the same cross correlation analysis on the jumbled spike train which has the same ISI distribution. See Supplementary Fig. 1 for examples of system functions obtained from single units utilizing this approach.

For the EEG data, the response in each channel was cross-correlated with the binaural rem-seq giving a multi-channel binaural temporal response function (mcBTRF) for each participant. The mcBTRF is then averaged across participants. In the mcBTRF, the response can be seen in several channels with different magnitudes and with different polarities. This is expected considering that the response from any given neural source will contribute voltage fluctuations in different EEG channels with different amplitudes and polarities depending on the geometrical configuration of the source relative to the sensors. To extract the dominant source and its spatial topography, we used principal component analysis (PCA) on the mcBTRF. The PCA implementation from `sci-kit learn` in Python was utilized. PCA was done in the time range of 0 to 500 ms and only the first principal component was taken as the source binaural temporal response function, sBTRF. The IAC sBTRF explained 94% of the variance of the mcBTRF, and the ITD sBTRF explained 70% of the variance. The PCA operation on the mcBTRF can be thought of as a spatial filter that generates a linear combination of the 32 channels such that the combination (i.e., the sBTRF) accounts for most of the variance in the mcBTRF. The scalp topographic map of the PCA weights to get the sBTRF can provide information about physical location of the dominant source tracking the binaural modulations. To estimate the variance of the sBTRF, we used a jackknifing (leave-one-out) procedure. The frequency response of the sBTRF was computed using the `freqz` function from the `scipy` library in python. The group delay of the sBTRF was computed by taking the slope of a fitted line of the unwrapped phase between 2.5 and 6 Hz. We choose 2.5 and 6 Hz because the sBTRF had high SNR in that region for both ITD and IAC. Noise floors for the EEG data were constructed by multiplying a randomly chosen half of the trials by -1 and then carrying out the same analysis as used to obtain the mcBTRFs. Ten noise-floor estimates were generated for each participant for each channel. The PCA weights from the actual mcBTRF were used on the noise-floor estimates to obtain the noise-floor estimates for the sBTRFs. See Supplementary Fig. 2 for a visual depiction for going from the 32-channel evoked response to the sBTRF.

Psychoacoustic experiments. We performed four psychoacoustic experiments, laid out in Fig. 1. All stimuli were generated using custom MATLAB scripts and delivered through the acoustic apparatus described previously. The FM tracking psychoacoustic experiment was done through our online platform that has previously been validated to yield average absolute thresholds that are comparable to lab-based data⁵³. For every task, participants were first given a brief demo to understand the task. For all binaural tasks, we modulated IAC dynamically instead of ITD. The reason we studied dynamic IAC and not dynamic ITD behaviorally is because with dynamic ITD, it is easy to introduce monaural artifacts. In fact, it is mathematically impossible to have a purely dynamic ITD. It is

quite easy for an unintended amplitude modulation to be introduced monaurally. Also, in part for the reason just discussed, much of the literature to this point has studied dynamic IAC instead of ITD, so to compare and put in context with previous work, using dynamic IAC measures made sense. Recent work has used dynamic ITD stimuli, but they also note the spectral fluctuations that become audible monaurally at higher modulation frequencies¹⁶.

Perceptual limits for detecting binaural modulations. To measure human ability to detect binaural modulations, we used the method of constant stimuli with the oscillating-correlation (OSCOR) stimulus. The OSCOR stimulus consists of noise tokens with sinusoidally varying IAC, and has been used previously in both behavioral and physiological studies^{14,15,20}. Each trial was 3-interval 3-alternatives-forced-choice with the target interval containing the OSCOR stimulus and the other two intervals containing interaurally uncorrelated noise (IAC = 0). We evaluated performance at octave frequencies between 5–320 Hz with 20 trials at each frequency. The OSCOR stimulus was band-limited between 0.2–1.5 kHz because of data suggesting fine-structure-based binaural cues may not be useful beyond 1.5 kHz³⁴. However, we repeated this experiment in one subject with white noise due to physiological data indicating cells with higher center frequencies can encode the fast OSCORs¹⁴. Indeed, one possibility is that fine-structure-based binaural cues may be detected for higher (beyond 1.5 kHz) carriers but that these cues don't inform spatial perception. The results of the measurement in the one participant with OSCOR applied to bandlimited (0.2–1.5 kHz) and to white noise (extending up to half the sampling rate) are shown in Supplementary Fig. 3. Because we wanted to simply demonstrate the white noise OSCOR has higher detection limit than the bandlimited OSCOR, we measured this in only one participant rather than several.

Limits for perceiving dynamic space. Several studies have anecdotally reported that with the OSCOR stimulus and other dynamic binaural stimuli, the perception of the stimulus appears to change from a spatialized image (i.e., moving in space) to a flutter around 6–10 Hz^{15,16,21}. We hypothesized that this switch would align with cortical temporal coding limits. Accordingly, we formally measured this switch in 1 participant using the method of limits with the OSCOR stimulus to formally measure this often anecdotally reported measure in the literature. There were 10 ascending and descending trials that started randomly between 3–6 Hz or 16–19 Hz respectively. The participant pushed a button indicating whether the perception of the stimulus had changed or not (either spatial to flutter or flutter to spatial) in each trial. If the perception had not changed, the frequency was increased (ascending trials) or decreased (descending trials) by 1 Hz until the change was noted.

Perceptual dynamics of spatial unmasking & comparison to physiology. The third behavioral task probed dynamic binaural unmasking, and was based on a previously published paradigm²². In this task the noise is uncorrelated (IAC = 0) except for a window of time in the middle of the stimulus where the noise becomes completely correlated (IAC = 1), see Fig. 4. While the noise is completely correlated, an anti-correlated (IAC = -1) 850 Hz tone, 20 ms in duration, is played coincidentally with correlated noise. The difference in IAC between the tone and the noise (i.e., the “N0S π ” configuration of the mixture) can be used to improve detection of the tone, i.e. a spatial unmasking effect. We varied the duration of the completely correlated period of the noise and measured detection thresholds for the tone using an

adaptive 2-up-1-down paradigm. The window durations we evaluated were 0, 50, 75, 100, 125, 150, 200, 400, 800, and 1600 ms. One participant quit during the last window duration, which was 800 ms, due to exhaustion, so that one data point was thrown out. Culling and Summerfield²² used this task to estimate the underlying binaural temporal analysis window by comparing the unmasking function (dB masking release vs. window duration function) with levels of unmasking that different window shapes would predict. Here, we measured the binaural temporal window physiologically using EEG. Thus, instead of fitting arbitrary window shapes, we analyze how well the physiologically measured temporal window, the sBTRF, quantitatively explains the entire behaviorally measured unmasking function. This was done in two steps. First, the sBTRF (normalized and shifted to sum to 1 and take non-negative values) was convolved with the background noise, and the maximum “internal” IAC of the noise is estimated in the window of overlap with the tone. Then a binaural masking level difference (BMLD), or detection improvement relative to a window duration of 0 is estimated from the known relationship between static IAC and BMLD⁵⁴, which is captured in Equation (1) below. van der Heijden and Trahiotis⁵⁴ found that this equation could account for 98% of the variance of behavioral BMLD data from Robinson and Jeffress¹³. Here, T_{No} is the mean threshold at the largest window size (1600 ms) and T_{Nu} is the mean threshold with no window present.

$$BMLD = -10 \log_{10} \left[(1 - IAC) + (IAC) \frac{T_{No}}{T_{Nu}} \right] \quad (1)$$

FM phase difference detection using web-based psychoacoustics. In response to the COVID19 pandemic, we developed and validated a web-based platform for conducting suprathreshold psychoacoustics experiments⁵³. We recruited 14 participants from Prolific in the 18–55 year age range. Each participant passed a headphone-use screening test, and a screening for normal hearing based on a suprathreshold speech-in-babble paradigm⁵³ before participating in the main FM experiment. One of the authors also completed the task, yielding a total of 15 total participants.

In the main task, participants were instructed to detect the difference between two frequency modulations at a given modulation rate, but applied to spectrally distant carriers. One carrier was always between 500–750 Hz, and the other carrier was chosen to be two octaves higher than the first. The modulation depth of the FM was 10% of the carrier frequency. The FM rates we evaluated were 4, 8, 16, 32, and 64 Hz and the phase difference between the FMs were 30, 60, 90, or 180 degrees. An example of the FM phase difference detection stimulus is shown in Supplementary Fig. 4. The stimulus duration was 1.5 s and had a sampling rate of 44,100 Hz. To eliminate potential onset effects in detecting the phase difference between the two FMs a discrete prolate-spheroidal sequence (DPSS) window was used to apply a 125 ms ramp, and the starting phase of the FMs in each interval was randomized. Each trial was organized in a 3-interval 3-AFC format, with non-target stimulus intervals containing in-phase FMs and the target interval containing the FMs with a phase difference. Mean and standard error parameters for detection accuracy were estimated using the median, and the median absolute deviation, respectively.

Statistics and reproducibility. Statistical inferences were drawn by looking at the distribution of data computed by the standard error of the mean and by directly comparing physiology to behavior using mean squared error. For physiological responses, we computed the standard error by using a jackknife approach.

The 95 % confidence interval computed from the standard error was used to evaluate where the physiologic response is clearly greater than the noise floor. All code to replicate this analysis is made available in our code availability statement. All raw data and code to generate the stimuli to generate these physiological responses is also made available through our data and code availability statements. The binaural behavioral responses estimated the standard error of the mean by dividing the standard deviation of the data by the square root of the number of samples. For the FM behavioral data, we utilized median absolute deviation to estimate the standard error of the mean because the online data had a tendency to have more potential outliers. Lastly, to compare the behavior and physiology, we utilized mean square error (MSE) which is a straightforward technique to measure the similarity of two signals. All code to replicate our analysis, generate all figures in this paper, and generate our stimuli is made available in our code and data availability statement to enhance reproducibility. All raw data are also made available as described in our data availability statement.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data is openly accessible and archived using Zenodo⁵⁵.

Code availability

The code to reproduce the figures in this work is publicly available on GitHub (<https://github.com/Ravinderjit-S/DynamicBinauralProcessing>) and archived using Zenodo⁵⁶. This repository also contains code for generating the stimuli used in this paper.

Received: 6 January 2022; Accepted: 15 September 2023;

Published online: 26 September 2023

References

- Joris, P. X., Sschreiner, C. E. & Rees, A. Neural processing of amplitude-modulated sounds. *Physiol. Rev.* **84**, 541–577 (2004).
- Liang, L., Lu, T. & Wang, X. Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J. Neurophysiol.* **87**, 2237–2261 (2002).
- Bendor, D. & Wang, X. Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J. Neurophysiol.* **100**, 888–906 (2008).
- Walker, K. M. M., Ahmed, B. & Schnupp, J. W. H. Linking cortical spike pattern codes to auditory perception. *J. Cogn. Neurosci.* **20**, 135–152 (2008).
- Schnupp, J. W. H., Hall, T. M., Kokelaar, R. F. & Ahmed, B. Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex. *J. Neurosci.* **26**, 4785 LP – 4795 (2006).
- Phillips, D. P., Hall, S. E. & Boehnke, S. E. Central auditory onset responses, and temporal asymmetries in auditory perception. *Hear. Res.* **167**, 192–205 (2002).
- Heil, P. First-spike latency of auditory neurons revisited. *Curr. Opin. Neurobiol.* **14**, 461–467 (2004).
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J. & Shamma, S. A. Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron* **61**, 317–329 (2009).
- Bendor, D. & Wang, X. Differential neural coding of acoustic flutter within primate auditory cortex. *Nat. Neurosci.* **10**, 763–771 (2007).
- Culling, J. F. & Colburn, H. S. Binaural sluggishness in the perception of tone sequences and speech in noise. *J. Acoust. Soc. Am.* **107**, 517–527 (2000).
- Hawley, M. L., Litovsky, R. Y. & Culling, J. F. The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* **115**, 833–843 (2004).
- Lingner, A., Grothe, B., Wiegrebe, L. & Ewert, S. D. Binaural glimpses at the cocktail party? *J. Assoc. Res. Otolaryngol.* **17**, 461–473 (2016).

13. Robinson, D. E. & Jeffress, L. A. Effect of varying the interaural noise correlation on the detectability of tonal signals. *J. Acoust. Soc. Am.* **35**, 1947–1952 (1963).
14. Joris, P. X., van de Sande, B., Recio-Spinoso, A. & van der Heijden, M. Auditory midbrain and nerve responses to sinusoidal variations in interaural correlation. *J. Neurosci. : Off. J. Soc. Neurosci.* **26**, 279–289 (2006).
15. Siveke, I., Ewert, S. D., Grothe, B. & Wiegand, L. Psychophysical and physiological evidence for fast binaural processing. *J. Neurosci.* **28**, 2043–2052 (2008).
16. Zuk, N. & Delgutte, B. Neural coding of time-varying interaural time differences and time-varying amplitude in the inferior colliculus. *J. Neurophysiol.* **118**, 544–563 (2017).
17. Joris, P. X. Neural binaural sensitivity at high sound speeds: single cell responses in cat midbrain to fast-changing interaural time differences of broadband sounds. *J. Acoust. Soc. Am.* **145**, EL45–EL51 (2019).
18. Fitzpatrick, D. C., Roberts, J. M., Kuwada, S., Kim, D. O. & Filipovic, B. Processing temporal modulations in binaural and monaural auditory stimuli by neurons in the inferior colliculus and auditory cortex. *J. Assoc. Res. Otolaryngol.* **10**, 579 (2009).
19. Scott, B. H., Malone, B. J. & Semple, M. N. Representation of dynamic interaural phase difference in auditory cortex of awake rhesus macaques. *J. Neurophysiol.* **101**, 1781–1799 (2009).
20. Grantham, D. W. Detectability of time-varying interaural correlation in narrow-band noise stimuli. *J. Acoust. Soc. Am.* **72**, 1178–1184 (1982).
21. Grantham, D. W. & Wightman, F. L. Detectability of varying interaural temporal differences. *J. Acoust. Soc. Am.* **63**, 511–523 (1978).
22. Culling, J. F. & Summerfield, Q. Measurements of the binaural temporal window using a detection task. *J. Acoust. Soc. Am.* **103**, 3540–3553 (1998).
23. Grantham, D. W. & Wightman, F. L. Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. *J. Acoust. Soc. Am.* **65**, 1509–1517 (1979).
24. Kollmeier, B. & Gilkey, R. H. Binaural forward and backward masking: evidence for sluggishness in binaural detection. *J. Acoust. Soc. Am.* **87**, 1709–1719 (1990).
25. Shackleton, T. M. & Bowsher, J. M. Binaural effects of the temporal variation of a masking noise upon the detection thresholds of tone pulses. *Acta Acust. U. Acust.* **69**, 218–225 (1989).
26. Scherg, M. & Von Cramon, D. Evoked dipole source potentials of the human auditory cortex. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials Sect.* **65**, 344–360 (1986).
27. Cohen, D. & Cuffin, B. Demonstration of useful differences between magnetoencephalogram and electroencephalogram. *Electroencephalogr. Clin. Neurophysiol.* **56**, 38–51 (1983).
28. Lewald, J. & Getzmann, S. When and where of auditory spatial processing in cortex: a novel approach using electrotomography. *PLoS One* **6**, e25146 (2011).
29. Vicić, D. et al. Human auditory belt areas specialized in sound recognition: a functional magnetic resonance imaging study. *NeuroReport* **17**, 1659–1662 (2006).
30. van der Zwaag, W., Gentile, G., Gruetter, R., Spierer, L. & Clarke, S. Where sound position influences sound object representations: A 7-T fMRI study. *NeuroImage* **54**, 1803–1811 (2011).
31. Stecker, G. C., Harrington, I. A. & Middlebrooks, J. C. Location coding by opponent neural populations in the auditory cortex. *PLoS Biol.* **3**, e78 (2005).
32. Salminen, N. H., May, P. J. C., Alku, P. & Tiitinen, H. A population rate code of auditory space in the human cortex. *PLoS One* **4**, e7600 (2009).
33. Magezi, D. A. & Krumbholz, K. Evidence for opponent-channel coding of interaural time differences in human auditory cortex. *J. Neurophysiol.* **104**, 1997–2007 (2010).
34. Brughera, A., Dunai, L. & Hartmann, W. M. Human interaural time difference thresholds for sine tones: the high-frequency limit. *J. Acoust. Soc. Am.* **133**, 2839–2855 (2013).
35. Verschooten, E. The upper frequency limit for the use of phase locking to code temporal fine structure in humans: a compilation of viewpoints. *Hear. Res.* **377**, 109–121 (2019).
36. Shackleton, T. M. & Palmer, A. R. The time course of binaural masking in the inferior colliculus of guinea pig does not account for binaural sluggishness. *J. Neurophysiol.* **104**, 189–199 (2010).
37. Zuk, N. J. & Delgutte, B. Neural coding and perception of auditory motion direction based on interaural time differences. *J. Neurophysiol.* **122**, 1821–1842 (2019).
38. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex (N. Y., NY)* **1**, 1–47 (1991).
39. Uusitalo, M. A. & Ilmoniemi, R. J. Space projection method for separating MEG or EEG into components. *Med. Biol. Eng. Comput.* **35**, 135–140 (1997).
40. Chintanpalli, A. & Heinz, M. G. Effect of auditory-nerve response variability on estimates of tuning curves. *J. Acoust. Soc. Am.* **122**, EL203–EL209 (2007).
41. Eggermont, J. J. Wiener and Volterra analyses applied to the auditory system. *Hear. Res.* **66**, 177–201 (1993).
42. Ringach, D. & Shapley, R. Reverse correlation in neurophysiology. *Cogn. Sci.* **28**, 147–166 (2004).
43. Henry, K. S., Sayles, M., Hickox, A. E., and Heinz, M. G. Divergent auditory-nerve encoding deficits between two common etiologies of sensorineural hearing loss. *J. Neurosci.* **39**, 6879–6887 (2019).
44. Recio-Spinoso, A., Temchin, A. N., van Dijk, P., Fan, Y.-H. & Ruggero, M. A. Wiener-kernel analysis of responses to noise of chinchilla auditory-nerve fibers. *J. Neurophysiol.* **93**, 3615–3634 (2005).
45. Wang, X. Neural coding strategies in auditory cortex. *Hear. Res.* **229**, 81–93 (2007).
46. Theunissen, F. E., Sen, K. & Doupe, A. J. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. *J. Neurosci.* **20**, 2315 LP – 2331 (2000).
47. Aptekar, J. W. Method and software for using m-sequences to characterize parallel components of higher-order visual tracking behavior in Drosophila. *Front. Neural Circuits* **8**, 1–15 (2014).
48. Chu, W. T. Impulse-response and reverberation-decay measurements made by using a periodic pseudorandom sequence. *Appl. Acoust.* **29**, 193–205 (1990).
49. Reid, R. C., Victor, J. D. & Shapley, R. M. The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Vis. Neurosci.* **14**, 1015–1027 (1997).
50. Shi, I. & Hecox, K. E. Nonlinear system identification by m-pulse sequences: application to brainstem auditory evoked responses. *IEEE Trans. Biomed. Eng.* **38**, 834–845 (1991).
51. Burkard, R. F., Finneran, J. J. & Mulsow, J. The effects of click rate on the auditory brainstem response of bottlenose dolphins. *J. Acoust. Soc. Am.* **141**, 3396–3406 (2017).
52. Harper, N. S. & McAlpine, D. Optimal neural population coding of an auditory spatial cue. *Nature* **430**, 682–686 (2004).
53. Mok, B. A. et al. Web-based psychoacoustics: Hearing screening, infrastructure, and validation. *Behav. Res. Methods* <https://doi.org/10.3758/s13428-023-02101-9> (2023).
54. van der Heijden, M. & Trahiotis, C. A new way to account for binaural detection as a function of interaural noise correlation. *J. Acoust. Soc. Am.* **101**, 1019–1022 (1997).
55. Singh, R. & Bharadwaj, H. *Dynamic Binaural Processing (sBTRF) Data*. <https://zenodo.org/record/5778003> (2021).
56. Singh, R. *Ravinderjit-S/DynamicBinauralProcessing: Dynamic Binaural Processing (sBTRF) Code*. <https://zenodo.org/record/8098476> (2021).

Acknowledgements

We would like to acknowledge Mark Sayles for his contributions to the development and execution of experiments in chinchillas to study brainstem level encoding of dynamic binaural cues. Mark Sayles also contributed to the development of the novel binaural systems identification approach used in this work. This work was supported by National Institutes of Health (NIH) grants R01DC015989 (HMB), and T32DC016853 (RS).

Author contributions

H.M.B and R.S. designed the study and wrote the manuscript. R.S. collected and analyzed the data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05361-5>.

Correspondence and requests for materials should be addressed to Hari M. Bharadwaj.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Jacqueline Gottlieb, Caitlin Karniski and Karli Montague-Cardoso.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023