Received: 1 May 2024 | Revised: 15 October 2024 | Accepted: 26 November 2024

DOI: 10.1111/ejn.16637

TECHNICAL NOTE

EJN European Journal of Neuroscience FENS WILEY

Measuring self-similarity in empirical signals to understand musical beat perception

Tomas Lenc[1,2] | **Cédric Lenoir**[1] | **Peter E. Keller**[3,4] | **Rainer Polak**[5,6] | **Dounia Mulders**[1,7,8,9] | **Sylvie Nozaradan**[1,10]

[1]Institute of Neuroscience (IONS), UCLouvain, Brussels, Belgium

[2]Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastian, Spain

[3]MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, Australia

[4]Center for Music in the Brain & Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

[5]RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Oslo, Norway

[6]Department of Musicology, University of Oslo, Oslo, Norway

[7]Computational and Biological Learning Unit, Department of Engineering, University of Cambridge, Cambridge, UK

[8]Institute for Information and Communication Technologies, Electronics and Applied Mathematics, UCLouvain, Louvain-la-Neuve, Belgium

[9]Department of Brain and Cognitive Sciences and McGovern Institute, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA

[10]International Laboratory for Brain, Music and Sound Research (BRAMS), Montreal, Canada

Correspondence
Tomas Lenc, The Institute of Neuroscience (IONS), Université catholique de Louvain (UCLouvain), 53 Avenue Mounier, B1200 Brussels, Belgium.
Email: tomas.lenc@uclouvain.be

Funding information
Fonds De La Recherche Scientifique - FNRS : FC17797; H2020 European Research Council, Grant/Award Number: 801872; HORIZON EUROPE Marie Sklodowska-Curie Actions, Grant/Award Number: 101148958; Danish National Research Foundation, Grant/Award Number: DNRF117

Edited by: Ali Mazaheri

Abstract

Experiencing music often entails the perception of a periodic beat. Despite being a widespread phenomenon across cultures, the nature and neural underpinnings of beat perception remain largely unknown. In the last decade, there has been a growing interest in developing methods to probe these processes, particularly to measure the extent to which beat-related information is contained in behavioral and neural responses. Here, we propose a theoretical framework and practical implementation of an analytic approach to capture beat-related periodicity in empirical signals using frequency-tagging. We highlight its sensitivity in measuring the extent to which the periodicity of a perceived beat is represented in a range of continuous time-varying signals with minimal assumptions. We also discuss a limitation of this approach with respect to its specificity when restricted to measuring beat-related periodicity only from the magnitude spectrum of a signal and introduce a novel extension of the approach based on autocorrelation to overcome this issue. We test the new autocorrelation-based method using simulated signals and by re-analyzing previously published data and show how it can be used to process

Abbreviations: ACF, autocorrelation function; DFT, discrete Fourier transform; ERP, event-related potential; EEG, electroencephalography; SNR, signal-to-noise ratio.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2025 The Author(s). *European Journal of Neuroscience* published by Federation of European Neuroscience Societies and John Wiley & Sons Ltd.

measurements of brain activity as captured with surface EEG in adults and infants in response to rhythmic inputs. Taken together, the theoretical framework and related methodological advances confirm and elaborate the frequency-tagging approach as a promising window into the processes underlying beat perception and, more generally, temporally coordinated behaviors.

KEYWORDS
autocorrelation, EEG, frequency-tagging, periodicity, rhythm

1 | INTRODUCTION

Humans across cultures show a remarkable capacity to coordinate movement in time with music. Such musical behaviors are often guided by a specific kind of internal time representation, typically referred to as the "beat" (Honing & Bouwer, 2018; Jones & Mcauley, 2005; Large, 2008; London, 2012). More specifically, the current paper focuses on instances where the term "beat" refers to an *internal representation consisting of seamlessly recurring periods that can be used to time movement*, and are elicited by, and temporally coordinated with, a rhythmic sensory stimulus. In fact, rhythms typically used in musical contexts often elicit an internal representation of a set of faster and slower beat layers that form a nested temporal structure called "meter" (Brochard et al., 2003; Cohn, 2020; Large et al., 2002; Lenc et al., 2021; Repp, 2008; Toiviainen et al., 2010).

What remains poorly understood is how the brain establishes a stable representation of a beat from complex sensory inputs such as music. This question is far from trivial since the periodic beats perceived when listening to rhythmic stimuli are almost never directly present in the acoustic input (Honing & Bouwer, 2018; London, 2012). In other words, rhythms that elicit perception of highly regular, often isochronous beats are themselves rarely regular to the same degree, not to speak of isochronous; instead, such musical rhythms mostly show various, sometimes only weak degrees of acoustic congruency with the beats they afford (Lenc et al., 2021; London et al., 2017). This illustrates a core property of beat representation, that is, its high degree of *invariance* with respect to the sensory input (Lenc et al., 2021; Nozaradan, Keller, et al., 2017). Hence, beat perception can be viewed as a form of *perceptual categorization* (Clarke, 1987; Goldstone et al., 2018; Holt & Lotto, 2010; Schulze, 1989; Windsor, 1993), that is, a function that maps an internal periodic template onto sensory inputs spanning a range of physical properties, including various arrangements of sounds in time (see Figure 1a) (Chemin et al., 2014; Large et al., 2015; Nozaradan et al., 2012; Phillips-Silver & Trainor, 2005; Su & Pöppel, 2012; Tal et al., 2017; Witek et al., 2014).

In this paper, we discuss how our understanding of beat perception can be advanced by capturing representations of beat periodicities reflected in behavioral and neural responses to rhythmic sensory inputs. We argue that beat periodicity can be reliably captured by measuring self-similarity of signals over time, and we focus on a practical implementation of this concept using a methodology based on the frequency-tagging approach. We develop an extension of frequency-tagging based on autocorrelation and propose tools to facilitate the practical application of this method to empirical measurements.

2 | CONCEPTUALIZING BEAT PERIODICITY AS SELF-SIMILARITY

A particular beat periodicity can be measured from behavioral or neural responses elicited by a rhythmic stimulus. By definition, any time-varying *response* (e.g. firing rate of a neuron, state of a population of neurons or spatial position of an effector) can be considered to reflect a particular internally represented beat when this response systematically repeats itself at the rate given by the beat period (see Figure 1b). In other words, a time-varying behavioral or neural response contains beat-related information if its dynamics exhibit *periodic self-similarity or recurrence* precisely locked onto the beat period (for an extended discussion, see Lenc et al., 2021).

This conceptualization can be framed in a broader context of systems neuroscience, particularly its central tenet that an internal representation of a particular variable is reflected in a *systematic relationship* between that variable and some attribute of the response measured from the system (Kriegeskorte & Wei, 2021; Rossion et al., 2023). In other words, the system is expected to generate selective (i.e. discriminant) responses to different values of the variable and reproducible (i.e. generalizable) responses to the same values of the variable across a wide range of conditions. Beyond this systematic pattern of similarities/dissimilarities based on

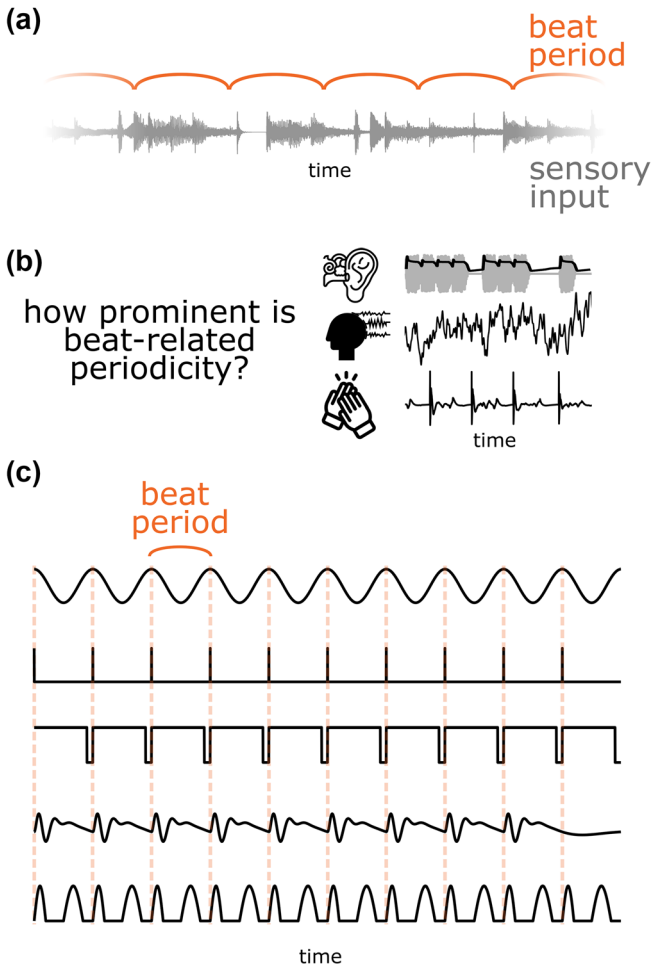


FIGURE 1 Beat as perceptual categorization. (a) Illustration of the categorization or mapping of external auditory input (in gray) onto an internal representation consisting of a periodic beat (in orange). The beat is represented abstractly as a series of regularly recurring periods. (b) A schematic illustrating that beat-related information, namely prominence of perceived beat periodicity, can be captured from various time-varying signals, including firing-rate dynamics of subcortical auditory neurons (top), electrophysiological brain activity (middle) or overt movement such as hand clapping (bottom). (c) A range of signals with different shapes but comparable periodic recurrence (i.e. self-similarity) at the rate of the beat. Each row shows an example signal in black, and the period of the beat is indicated by the orange curve on the top as well as the vertical dashed lines. Icon sources: "Ear" by Eucalyp, "EEG" by Aenne Brielmann and "clapping" by Adrien Coquet from the Noun Project under CC BY 3.0 license.

the targeted functional phenomenon (also referred to as a "second-order isomorphism") (Shepard & Chipman, 1970), the particular form of the response can be considered irrelevant. Approaches relying on second-order isomorphisms are increasingly popular in systems neuroscience, as they offer a comparable method to capture internal representations from a range of response modalities with minimum number of assumptions (Kriegeskorte et al., 2008; Kriegeskorte & Kievit, 2013).

Along these lines, we posit that the representation of a musical beat periodicity can be captured in the response signal based on the structure of self-similarity *over time*. In other words, a response representing a beat with a particular period should exhibit similar values at particular positions within the beat period across many repetitions of the period and not otherwise. The advantage of using such a self-similarity framework is that it allows *abstracting away from specific properties of the response that are not relevant to the phenomenon of interest*, thus focusing on beat operationalized only as periodic recurrence. Indeed, a particular periodic beat can be, in principle, equivalently reflected in response signals with a wide range of shapes and properties, as long as these signals show a high degree of periodic self-similarity (see Figure 1c).

3 | CAPTURING BEAT-RELATED INFORMATION FROM BODY MOVEMENT: FINGER-TAPPING

A widely used method in the research field to capture information about the perceived beat is based on measuring periodic recurrence in finger-tapping responses from participants explicitly instructed to tap the beat that they perceive when listening to a rhythmic stimulus (Large et al., 2015; McKinney & Moelants, 2006; Parncutt, 1994; Repp & Su, 2013; Toiviainen & Snyder, 2003). This specific form of behavioral response is assumed to represent beat-related information through a characteristic form of self-similarity, whereby the effector systematically impacts the surface at a fixed time position within each beat period and not otherwise. Hence, the main advantage of tapping is that it provides a well-defined marker of each beat period. The rate of the perceived beat can be therefore easily identified *directly* from the time intervals between successive taps. It should be noted that the self-similarity framework taken here does not need to assume that finger impact provides a marker of absolute beat position, i.e., the specific phase or alignment of the periodic beat with respect to the stimulus (Aschersleben, 2002; Repp, 2005; Repp & Su, 2013).

However, critical insights into the nature of beat perception may require capturing beat-related information in populations that may not be able to perform instructed behaviors such as tapping the perceived beat, for example patients, young infants, and non-human animals (Lenc et al., 2022; Merchant & Honing, 2014; Nozaradan, Schwartze, et al., 2017; Phillips-Silver & Trainor, 2005; Sifuentes-Ortega et al., 2022). In fact, a behavioral

outcome such as overt movement produced intentionally and following explicit instructions may be biased by decisional and cognitive factors, thus potentially leading to instances where one-to-one correspondence between the rate of movement and the perceived beat cannot be guaranteed (Lenc et al., 2021).

Moreover, the nature of the movement performed as an outcome of the explicit instruction may bias the response due to mechanical constraints. For example, hand clapping might express beats up to a much lower rate compared to finger tapping (Bamford et al., 2023; Repp & Su, 2013; Toiviainen et al., 2010); conversely, bimanual finger tapping might reflect beats at higher rates than unimanual tapping (Loseby et al., 2001). In addition, overt movement itself may significantly modulate processing of musical rhythm through concomitantly produced sensory input (e.g. proprioceptive) (Cannon, 2021; Manning & Schutz, 2013; Su & Pöppel, 2012). Therefore, movement-free approaches to capturing beat representation covertly may prove relevant to complementing existing research paradigms for advancing our understanding of the perceptual phenomenon, including its embodied nature and interactions with overt movement.

One way to go beyond the limitations of behavioral tasks is to measure beat-related information from neural responses to rhythmic stimuli. However, the dynamics of brain activity elicited by rhythmic stimuli does not provide well-defined temporal markers of beat periods, in contrast to finger tapping. That is, physiological responses (including brain activity) elicited by rhythmic sounds typically constitute complex waveforms including slow fluctuations rather than series of transient discrete events with clear temporal onsets (Bamford et al., 2023; Gámez et al., 2019; Leman & Naveda, 2010; Nozaradan et al., 2018). Consequently, time-domain methods typically used to capture periodic recurrence in finger tapping cannot be applied to neural responses. Instead, a number of recent studies have adopted a more general approach based on frequency-tagging, which allows beat-related information to be captured by quantifying periodic recurrence in a range of continuous time-varying signals.

4 | FREQUENCY-TAGGING BASED ON MAGNITUDE SPECTRUM ANALYSIS

Frequency-tagging is based on the fact that periodic recurrence in a signal is reflected in the frequency domain as peaks occurring at an exact set of frequencies corresponding to (1/recurrence period) and harmonics (Bach & Meigen, 1999; Norcia et al., 2015; Regan, 1989; Rossion et al., 2020). This can be illustrated with an example of a signal shown in Figure 2, consisting of 11 repetitions of an identical segment. The signal therefore contains periodic self-similarity at the repetition rate of the segment (i.e. at the rate corresponding to the inverse of segment duration). Taking the discrete Fourier transform (DFT) of this example signal yields a spectrum with narrow peaks positioned exactly at frequencies corresponding to integer multiples, i.e. harmonics, of the segment repetition rate. This is because *any* periodically recurring signal can be described as a weighted sum (a linear combination) of sine waves that complete an integer number of cycles within the signal repetition period. The magnitudes of peaks observed in the obtained spectrum are thus proportional to these weights.

In fact, if the signal *perfectly* repeats itself with a fixed period (and an exact integer number of these repetitions has been captured), no other frequencies besides the

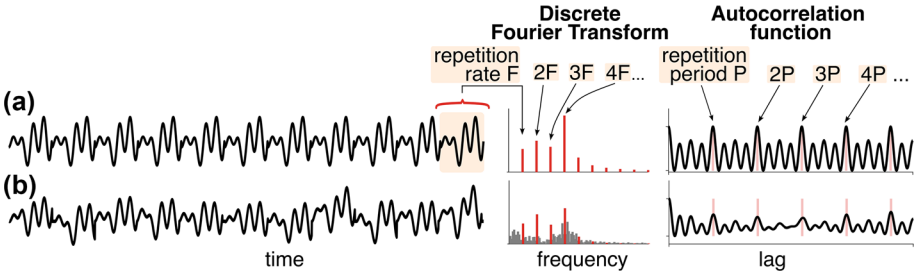


FIGURE 2 Periodic recurrence is reflected in the discrete Fourier transform (DFT) and autocorrelation function of the signal. (a) A signal was generated by repeating a smooth complex trajectory periodically every P seconds ($N = 11$ repetitions), giving rise to recurrence at the rate $1/P = F$ times per second. Taking the DFT of this signal yields a magnitude spectrum with narrow peaks centered at frequencies corresponding to integer multiples (i.e. harmonics) of F. Autocorrelation function of the signal reveals local maxima at lags corresponding to integer multiples of P. (b) Signal generated by randomly varying the shape as well as onset time (i.e. phase consistency) of the repeated trajectory, which decreases the self-similarity of the signal. The peaks at the harmonics of F become relatively less prominent in the spectrum, as do the peaks at the multiples of P in the autocorrelation function.

harmonics of the repetition rate are required to describe the signal. In contrast, reduced self-similarity of the repeated signal leads to lower peaks at the harmonic frequencies as the magnitudes become less concentrated and "leak" to other frequencies unrelated to the repetition rate. These principles are powerful because they can be harnessed to identify and quantify periodic recurrence by (1) "tagging" a set of objectively defined harmonic frequencies given a rate of interest and (2) measuring the prominence of peaks at these frequencies in the spectrum of any time-varying response.

4.1 | Taking harmonics into account

While observing peaks at a set of harmonics indicates the presence of periodic recurrence in the signal, the relative distribution of magnitudes across harmonics is determined by the specific shape of the recurring trajectory in the signal (see also Zhou et al., 2016). Consequently, to capture periodic recurrence *irrespective* of the shape of the repeating pattern (see Section 2), *harmonics must be considered together as a whole* (Retter et al., 2021).

It is important to underscore that the magnitude of a single harmonic, or generally of a single frequency (which can be seen as a very narrow band-pass filter whose width corresponds to the spectral resolution), can be essentially thought of as describing a correlation between the signal and a sine wave with that particular frequency (that is, the DFT at a given frequency answers the question: "How much does the signal resemble a sine wave with this frequency?"). While sine waves have mathematical properties that make them suitable for signal decomposition, periodic recurrence as discussed in the context of the current paper extends beyond sine waves. In other words, if the aim of frequency-tagging here is to capture the representation of a beat simply defined as a periodically recurring trajectory in the signal irrespective of its particular shape, there is no reason to assume that the trajectory must correspond to a sine wave.

4.2 | Optimizing the design of rhythmic stimuli for frequency-tagging

To precisely tag frequencies expected to capture beat-related information in the neural response, the *exact* period of the to-be-measured beat must be *specified* a priori (Bach & Meigen, 1999; Norcia et al., 2015). Therefore, frequency-tagging is typically used in tandem with another method, which allows identifying the beat period most consistently perceived by participants when listening to the same stimulus. For instance, participants can be asked to tap the perceived beat while listening to the rhythmic stimulus in a dedicated session (Lenc et al., 2018, 2020; Nozaradan et al., 2012, 2018). However, the period of the perceived beat estimated from tapping data will necessarily be subject to noise. Therefore, reducing the number of plausible beat periods to a few well-defined candidates is particularly critical.

To limit the number of plausible beat periods, frequency-tagging studies have often taken advantage of the common practice of using rhythmic sequences constructed by assigning sounds to positions corresponding to an evenly spaced (i.e. isochronous) grid of time points (Grahn & Brett, 2007; Nozaradan et al., 2012; Povel & Essens, 1985) (see example in Figure S1A). The advantage of this approach is that arranging the constituent sounds on an isochronous time grid restricts the number of possible beat periods that are likely to be induced by the stimulus. Specifically, grid-based rhythmic stimuli only allow beats with periods corresponding to integer multiples of the grid interval, to enable temporal coordination between the stimulus and the perceived beat. Therefore, knowing the grid interval used to construct the stimulus helps in precisely inferring the rate of the beat that the rhythmic input is susceptible to induce.

For example, if the stimulus was constructed on a .2-s interval grid, observing the average tapping rate of 1.22 Hz (∼.82-s inter-tap interval) in response to that stimulus indicates that the most likely rate of the perceived beat was 1.25 Hz, i.e. a .8-s period, corresponding to exactly four grid intervals. Then, informed by this observation, frequency-tagging can subsequently capture beat-related information by measuring magnitudes at harmonics of 1.25 Hz in neural responses to the same rhythm. It is important to note that the hypothesis about the to-be-measured beat period can be based on data from the same population yet may also come from a different population than the one producing the responses that are analyzed with frequency-tagging. This would presumably depend on the particular goal and context of each experiment, and as such is thus an empirical question. For example, tapping data from a group of healthy Western adults may be used to make a precise hypothesis about the period of a beat subsequently probed in neural responses to the exact same stimuli obtained from young human infants (Lenc et al., 2022), patients (Nozaradan, Mouraux, et al., 2016, Nozaradan, Schwartze, et al., 2017), or non-human animals (Rajendran et al., 2017, 2020).

In addition to constructing rhythmic stimuli based on an isochronous grid of time intervals, another way to reduce the number of plausible beat periods is to use rhythms made up of a repeating rhythmic pattern,

seamlessly looped to form a long sequence (Lenc et al., 2018; Nozaradan et al., 2012; Tal et al., 2017). Notably, using looped rhythmic patterns has high ecological validity, since periodically repeating rhythms are characteristic for music, particularly dance music, across many cultures (Butler, 2006; Câmara & Danielsen, 2018; London, 2012; London et al., 2017; Margulis, 2014; Witek, 2017).

The advantage of using repeated rhythms is that if the duration of the rhythmic pattern is not too long, humans tend to pick up the periodic recurrence from the pattern repetition (Asokan et al., 2021; Harrison et al., 2020; Norman-Haignere et al., 2022) and use it to constrain the perceived beat period such that the beat is *nested* within the pattern repetition cycle (Milne et al., 2023; Parncutt, 1994; Temperley & Bartlette, 2002). In other words, stimuli made of repeated rhythmic patterns tend to induce perception of beats that complete an integer number of periods within each rhythmic pattern repetition. Therefore, periodically recurring rhythmic patterns are optimal to ensure an integer number of perceived beat periods within the analyzed response signal (one of the assumptions of frequency-tagging). In addition, as discussed in the following section, looped rhythms offer a straightforward way to obtain a standardized estimate of self-similarity at the beat period, which can be directly compared across various signals.

4.3 | Standardization and lower level sensory confounds

Frequency-tagging can be used to measure how strongly a particular periodicity corresponding to a given beat is represented in a response signal. To this end, a set of harmonic frequencies of interest is selected based on a specific choice of the to-be-measured beat period (informed by behavior and stimulus parameters, as detailed in the above section). Hereafter, these harmonics will be referred to as "beat-related frequencies". Observing peaks centered at beat-related frequencies in the spectrum of the response indicates a recurring self-similarity at the rate of the beat.

However, such periodic recurrence can be only interpreted after taking the physical features of the stimulus into account. Specifically, it is critical to ensure that the physical features of the rhythmic input fluctuate over time in a way that is largely orthogonal to the periodicity of the perceived beat. The reason is that if the periodicity of the perceived beat matches a periodicity that is prominent in one of the time-varying physical features of the stimulus to which the neural system is susceptible to respond to, the elicited response will likely also show a high degree of self-similarity at this same periodicity. Yet, such a periodic response, while potentially related to internal representation of the beat, would be virtually impossible to dissociate from these lower level sensory confounds obtained at the same periodicity. Arguably, this issue is encountered in any instance of high-level perceptual categorization, i.e. where the internal representation is characterized, by definition, by a high degree of invariance from the physical features of the stimulus (see, for example, Rossion and Retter [2020] for a discussion in the context of face perception).

A simple way to illustrate the issue of lower level sensory confounds in the case of capturing beat representation is when the stimulus is an auditory metronome sequence, where a click sound occurs periodically. In a hypothetical experiment, one would first collect finger-tapping data to determine the rate of the perceived beat elicited by the metronome. If the metronome tempo is within a reasonable range, it is likely that participants will tap a beat at the rate of the metronome (Repp, 2005; Repp & Su, 2013). Based on the tapping, one may choose to use frequency-tagging to measure the representation of the perceived beat in brain activity elicited by the metronome during listening without movement. Observing narrow peaks centered at the metronome rate and harmonics in the spectrum of the neural response, that is, periodic recurrence at the rate of the perceived beat, may be interpreted as prominent neural representation of the beat. However, the neural response may be mostly explained by lower level sensory processing, whereby each metronome click evokes a consistent response from the auditory system, *regardless of whether these recurrent responses are related to an actual internal representation of the beat or not*. In other words, this instance does not provide a possibility to check whether the invariance with respect to the physical features of the stimulus—a criterion that is *central* to the definition of beat as a high-level perceptual categorization process—is met or not.

While the example of a metronomic stimulus seems trivial, it has important implications. In general, if the temporal arrangement of physical features in the stimulus itself shows prominent periodic recurrence at the rate of the beat, it is difficult to tease apart the relative contribution of (i) beat representation and (ii) tracking of stimulus features to the elicited neural and behavioral responses (Lenc et al., 2021; Nozaradan, Keller, et al., 2017). To go beyond lower level sensory confounds, one must show that the periodic recurrence observed in the response is above and beyond what could be expected if the system was merely driven by the physical features of the sensory input. A valid approach therefore needs to rely on comparing the relative prominence of beat-related periodicities between stimulus features and corresponding response.

However, frequency-tagging in its basic form does not offer a valid way to directly compare periodic recurrence in the stimulus and response. The reason is that stimulus features (e.g. sound intensity) and response (e.g. electrophysiological brain activity) have different units. Notably, while raw magnitudes at the beat-related frequencies are sensitive to changes in periodic recurrence, they are also affected by the unit and scale of the signal. For example, changing the overall gain of the signal without modifying its temporal structure leads to a proportional change of magnitudes across *all* frequencies in the spectrum. Consequently, solely measuring magnitudes at beat-related frequencies may be misleading in general, as the magnitudes will be affected by the scale of the signal.

Instead, the fact that magnitudes at all frequencies in the spectrum change proportionally as a function of signal scale may be exploited to obtain a standardized measure of magnitude at beat-related frequencies. Specifically, magnitudes at beat-related frequencies can be expressed *relative to magnitudes at other frequencies in the spectrum*. By expressing how much the beat-related frequencies stand out in the spectrum in comparison to other frequencies, this standardization thus provides a unitless measure that can be compared across stimulus and response, irrespective of the scale of these different signals. Moreover, taking other frequencies in the spectrum into account is important to quantify periodic recurrence. Indeed, measuring magnitude at beat-related frequencies is not enough to infer the prominence of self-similarity in the signal at the rate of the beat because one must also ensure that there are no prominent peaks at other, "beat-unrelated" frequencies in the spectrum.

How to select the "beat-unrelated" frequencies for standardization? A powerful approach is to select a set of frequencies that may be expected a priori based on the design of the stimulus but do not correspond to the harmonics of the to-be-measured beat rate. Such an approach is easily applied when the stimulus consists of a repeating rhythmic pattern. As the rhythmic pattern in the physical input periodically repeats R_P times per second (pattern repetition rate), the stimulus is expected to elicit a corresponding periodic response simply due to lower level sensory tracking.

Hence, the spectrum of the response is expected to contain peaks at frequencies corresponding to the *harmonics of the pattern repetition rate*, i.e., R_P, $2R_P$, $3R_P$, $4R_P$, etc. In fact, these harmonics will isolate any response as long as it consistently occurs at every repetition of the rhythmic pattern. This idea is powerful because it offers a straightforward approach to isolate a response consistently elicited by the rhythmic pattern from noise (Nozaradan et al., 2018; Sifuentes-Ortega et al., 2022). In other words, harmonics of the pattern repetition rate capture any response that is time-locked onto the rhythmic pattern in a similar way as averaging across many trials isolates activity consistently elicited by (i.e. time-locked to) a transient sensory input (Luck, 2014). Importantly, the *relative distribution* of magnitudes across these harmonics will depend on the shape of the response *within* the cycle of the repeated pattern. For example, if the response prominently represents a beat with a period three times faster than the pattern repetition rate, the response should comprise three repetitions of a consistent trajectory evenly spanning the cycle of the rhythmic stimulus. The more self-similar these nested repeated trajectories within the response, the more magnitude will be concentrated at frequencies corresponding to the beat-related frequencies at $3R_P$, $6R_P$, $9R_P$, etc. *relative to the other harmonics of the pattern repetition rate* (see Figure 3).

In order to obtain a standardized measure of this relative prominence, the magnitudes at beat-related frequencies can be standardized by considering a set of beat-unrelated frequencies corresponding to non-overlapping harmonics of the pattern repetition rate (i.e. R_P, $2R_P$, $4R_P$, $5R_P$, $7R_P$, etc.). The standardized relative prominence of the beat-related frequencies can be then quantified by *z*-scoring magnitudes across all frequencies of interest (beat-related and beat-unrelated frequencies) and taking mean *z*-score at beat-related frequencies.

In the current paper, *z*-scoring will be used for standardization, since the measure has several relevant properties. In particular, *z*-scoring is robust in situations where the analyzed set of values at beat-related and -unrelated frequencies may (i) have different unit or scale, (ii) be shifted by an offset and (iii) contain positive and negative values. Having a measure of periodic recurrence invariant to these signal characteristics is important in the context of frequency-tagging, especially when a noise-subtraction procedure is used to process empirical measurements (see "Methods" and Figure S4). Likewise, these properties of *z*-score standardization will prove particularly relevant when interpreting signal autocorrelation, as discussed in detail in Section 5 (see also Figures 6 and S8).

Despite the relevant characteristics of the *z*-score standardization, it is important to acknowledge its potential undesired behavior in certain extreme cases. For instance, if the analyzed values are near constant, any small deviations due to, for example, noise, may cause large changes in the *z*-scored values. This is because the *z*-score computation involves division by the standard deviation across all (beat-related end -unrelated) frequencies. Hence, if there is little variance across frequencies in

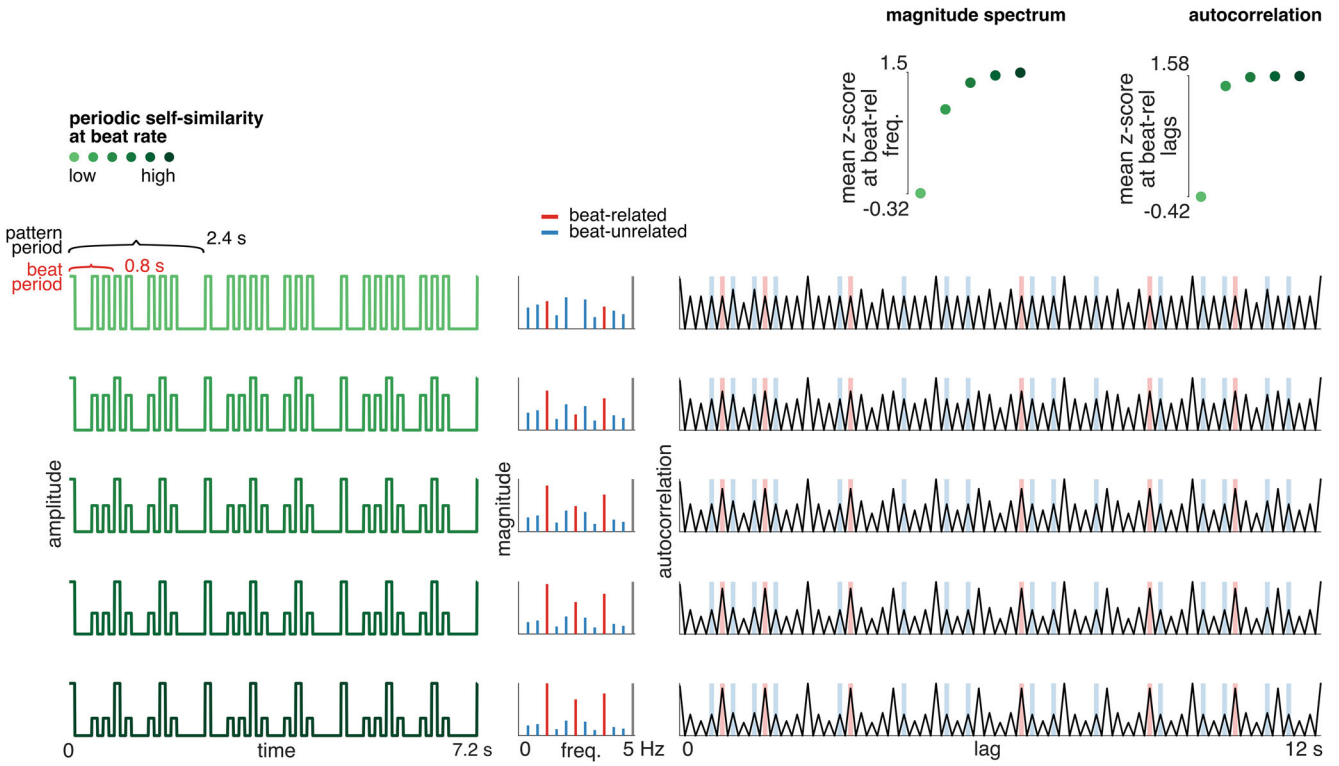


FIGURE 3 Sensitivity of frequency-tagging to periodic recurrence. Each row shows a simulated signal based on a repeating 2.4-s-long rhythmic pattern ($N = 20$ repetitions) and a square-wave unitary response. Across rows, the periodic recurrence at the rate of the beat (indicated by the red bracket on the top) is gradually increased (indicated by green color gradient). On the left, each row shows a 7.2-s-long segment of the signal in the time domain, corresponding to three full repetitions of the constituent rhythmic pattern (as indicated by the black bracket on the top). The magnitude spectrum of the signal is shown in the middle, and the corresponding autocorrelation function (trimmed at 12 s for visualization purposes) is shown on the right. Beat-related frequencies/lags are highlighted in red and beat-unrelated frequencies/lags in blue. The mean *z*-scored magnitude at beat-related frequencies and the mean *z*-scored autocorrelation at beat-related lags are shown on the top right of the figure. Each point corresponds to a single condition. Frequency-tagging based on both, the magnitude spectrum and autocorrelation, is sensitive to periodic recurrence, as indicated by a progressively increasing beat-related *z*-score as a function of periodic recurrence in the signal.

the response (relative to noise), the *z*-score may become too sensitive to even small changes in the measured values (and in some extreme cases, such as for an isochronous rhythm where each grid point contains a unitary event, even intractable). In fact, similar issues arise in any method that involves normalization (e.g. ratio or contrast), when the measured values are near zero (i.e. in the case of low signal-to-noise ratio [SNR]).

It is worth emphasizing that *z*-scoring as used here should not be taken as a statistical measure (related to Gaussian distribution) but only as a way to obtain a standardized measure that is comparable across signals of various origins and characteristics. Notably, there are several other possible standardization methods (e.g. calculating differences in percentage between mean values at beat-related and -unrelated frequencies). Taken together, it is important that an informed choice of a particular standardization method is made case by case, *based on the knowledge of the stimulus as well as the measured empirical signals*. Ideally, several methods could be used in a complementary way to confirm the convergence of experimental results (as has been done in many prior studies, e.g. Nozaradan, Mouraux, et al., 2016; Lenc et al., 2018, 2022).

As shown in Figure 3, mean *z*-score at beat-related frequencies captures the gradual transformation from a simulated response with little periodic recurrence at the rate of a to-be-measured beat towards a response where this recurrence is particularly prominent. In fact, this measure is sensitive to changes in periodic recurrence at the beat rate irrespective of the particular shape of the signal (see Figure 4), which is fundamental for capturing beat representation based on the structure of self-similarity over time (see Section 2). Critically, as the *z*-scored magnitude at beat-related frequencies is invariant to unit and scale, the measure can be calculated from the spectrum of a time-varying stimulus feature (e.g. sound intensity) and directly compared to the beat *z*-score calculated from a

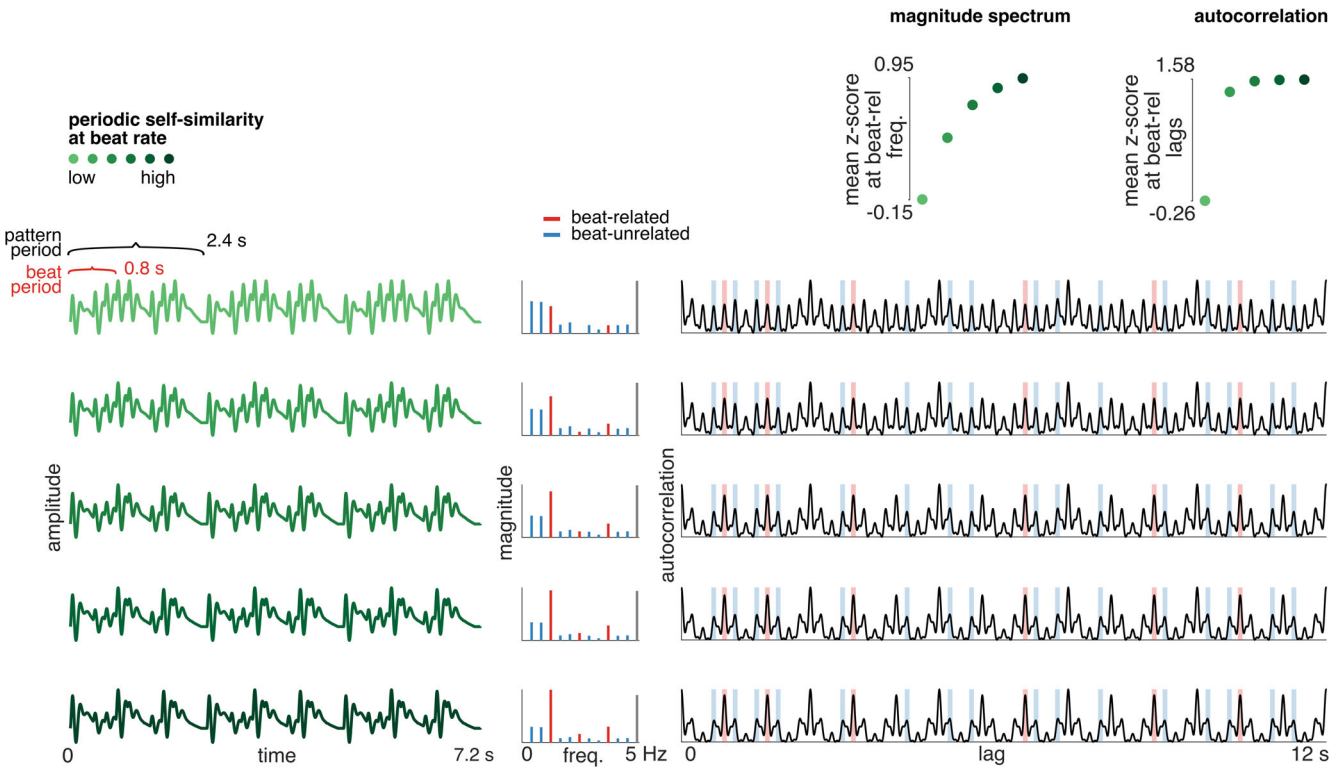


FIGURE 4 Sensitivity of frequency-tagging to periodic recurrence in signals comprising smooth complex waveforms. The structure is identical to Figure 3, but instead of a square-wave, a complex waveform is used as a unitary response.

corresponding response (e.g. neural activity). If the system selectively enhances the representation of the beat, beyond lower level sensory tracking of stimulus features, beat-related frequencies (as captured by their mean *z*-scored magnitude) should stand out in the spectrum of the response significantly more than in the spectrum of the stimulus (Nozaradan, Keller, et al., 2017).

However, recently, it has been suggested that a selective enhancement of beat periodicity as measured in the magnitude spectrum of neural responses may be driven by response properties that are unspecific to periodicity (Rajendran et al., 2017). Specifically, the relative prominence of beat-related frequencies may change depending on the shape of the signal, even when the periodic recurrence itself remains constant. This is illustrated in Figure 5, where a simulated rhythmic signal with a given amount of periodic recurrence at the rate of the beat is assembled from unitary events corresponding to a simple square wave. Depending on duration of the square-wave event, the mean *z*-scored magnitude at beat-related frequencies changes, despite the fact that the periodic recurrence (i.e. self-similarity) within the signal at the rate of the beat remains constant. To address this issue, we propose a complementary method based on autocorrelation, which can be considered an extension within the frequency-tagging approach.

5 | AUTOCORRELATION-BASED ANALYSIS

The goal of the novel autocorrelation-based extension is to obtain a measure of periodic recurrence that is *invariant* to the particular shape of the recurring signal. The autocorrelation approach rests on the fact that *a signal with strong recurrence at a particular period will be highly correlated with a time-lagged version of itself, as long as the time lag corresponds to an integer multiple of that period* (for examples of autocorrelation being used in the context of beat processing, see Brown, 1993; Tzanetakis & Cook, 2002; Toiviainen & Eerola, 2006; Ravignani & Norton, 2017).

One way to capture this self-similarity could be to compute the correlation between the original and time-shifted version of the signal, i.e. lagged exactly by one period of interest. The advantage of correlation methods, such as Pearson's correlation, is the fact that it is normalized, hence invariant to offset and scale of the signal (see Figure 6a,b). However, as illustrated in Figure 6c, the Pearson's correlation coefficient is affected by changes in the shape of the signal that do not impact periodic recurrence.

To achieve invariance with respect to the signal shape, the correlation computed after time-shifting the

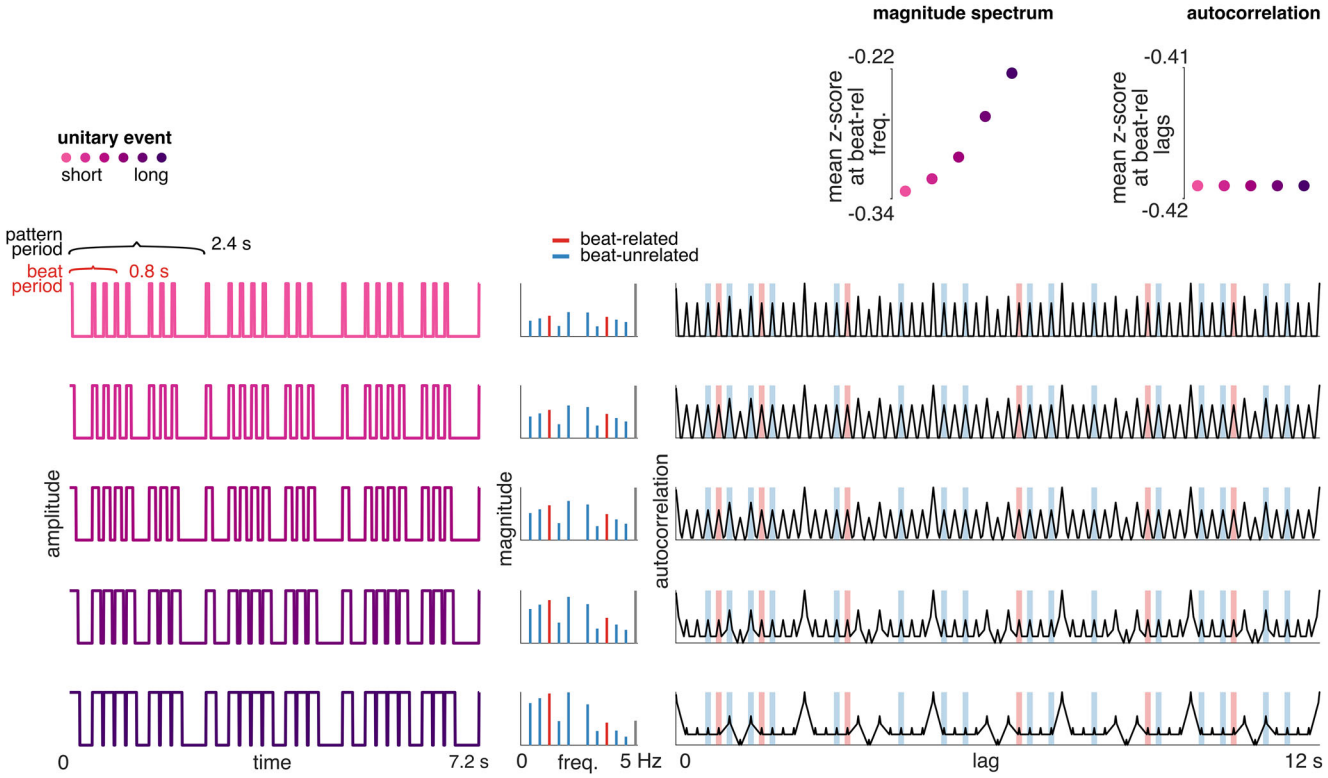


FIGURE 5 Effect of unitary response shape on the estimated periodic recurrence. Same as Figure 3, but periodic recurrence at the beat rate was kept constant across conditions. Instead, the shape of the unitary response making up the signal was gradually changed across conditions (shown in separate rows and indicated by the color gradient). Specifically, the shape was gradually changed from short duty cycle to long duty cycle (duty cycles referring to the ratio between the time interval where the unitary response shows a non-zero value and the inter-onset interval determined by the periodic grid of time intervals used to construct the rhythmic stimulus). Despite the periodic recurrence of the signal being constant, the mean *z*-score at beat-related frequencies as captured from the magnitude spectrum changes across conditions (shown on the top right). On the other hand, the mean *z*-scored autocorrelation at beat-related lags is identical across conditions.

signal by the target lag (i.e. the lag corresponding to the to-be-measured beat period) must be expressed *relative* to correlation values based on other, non-target lags, as determined based on the evenly spaced grid of time intervals used to construct the rhythmic stimulus (see Section 5.1 below). In particular, the autocorrelation at the target lag can be standardized using *z*-scoring, while taking values at several other lags into account. This yields a measure generally invariant to shape changes that do not affect the periodicity of the signal (see Figures 6 and S8). *z*-score standardization can be applied directly to Pearson's correlation coefficients computed after shifting the signal by the different lags of interest.

The approach is likewise valid for values obtained from a non-normalized circular autocorrelation function, which can be efficiently calculated from the DFT of the signal (see Section 7). As discussed in the following sections, the ability to compute autocorrelation via DFT turns out to be critical when accounting for the effect of noise on the signal of interest (see Section 5.2).

5.1 | Selecting beat-related and -unrelated lags

Given its high specificity to periodic recurrence in signals, autocorrelation can be exploited to capture beat-related information in response signals elicited by rhythmic stimuli. Critically, the approach allows direct comparison of periodic self-similarity in the response and in the physical stimulus that elicited it, thus providing a way to avoid lower level sensory confounds. Moreover, it is important to stress that the autocorrelation-based extension relies, for the most part, on principles discussed above in the context of frequency-tagging studies using the magnitude spectrum of signals.

First, a specific decision must be made about the target beat period one wishes to measure in the response. As for magnitude spectrum-based frequency-tagging, this can be informed by the design of the stimulus sequence combined with behavioral responses. Subsequently, lags corresponding to integer multiples of the target beat

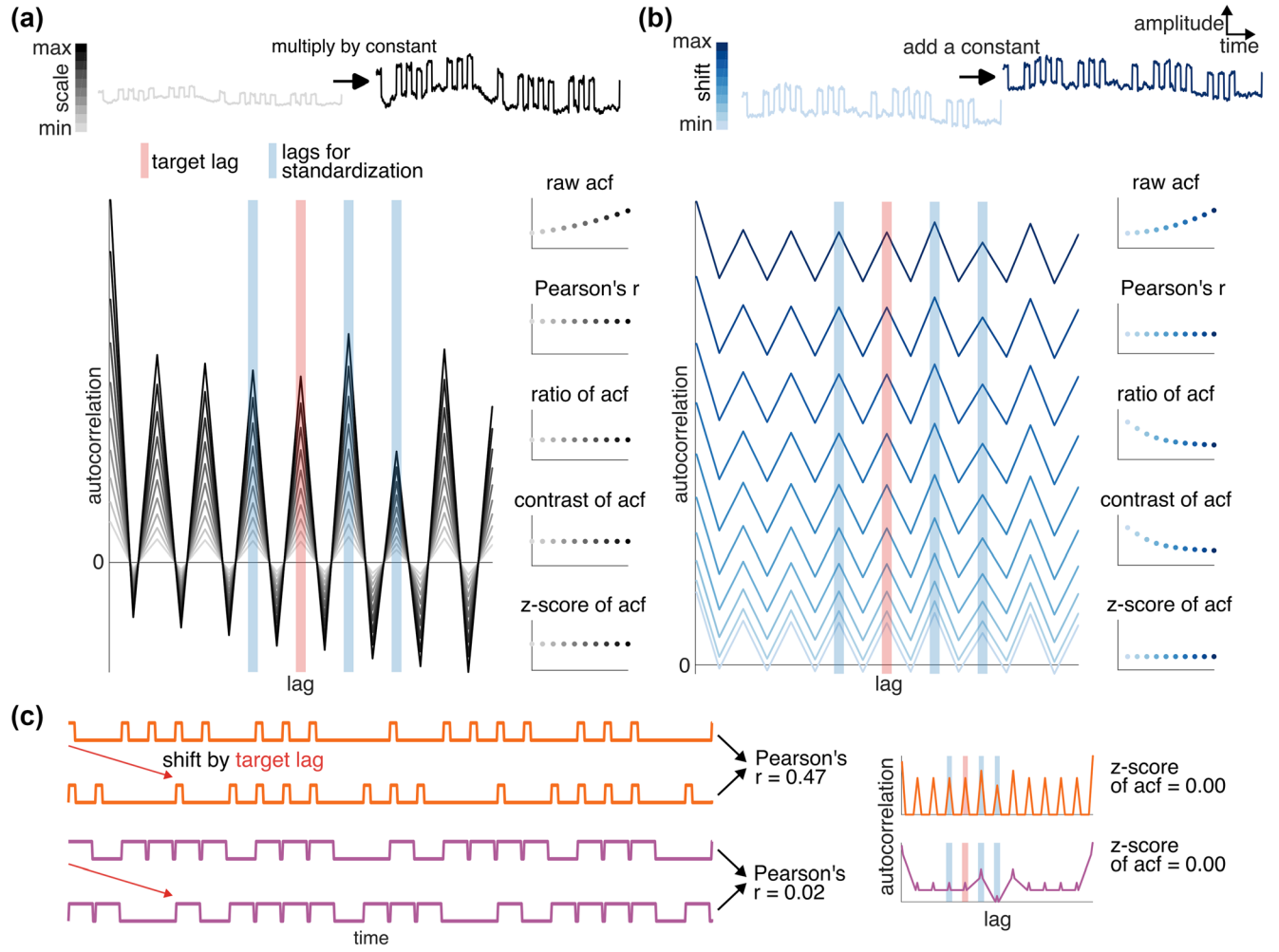


FIGURE 6 Invariance of the autocorrelation measure to signal features that do not affect periodic recurrence. (a) Sensitivity to scale of the signal. An identical signal (made up of 20 repetitions of a 2.4-s-long pattern) is scaled by a range of factors (indicated by color gradient), as illustrated on the top. The autocorrelation functions (ACF) of the scaled signals are shown on the bottom. The red vertical line indicates a target lag capturing a chosen period of recurrence. The blue vertical lines indicate several other lags used for standardization. The plots on the right show the autocorrelation at the target lag obtained (and standardized) using several different methods. Each point corresponds to one scale factor applied to the signal. "Raw ACF" value was directly extracted at the target lag from the ACF. "Pearson's *r*" was calculated using the time-domain signal circularly shifted by the target lag. Ratio, contrast, and *z*-score were calculated from ACF values at the target and standardization lags using equations provided in Section 7. The figure shows that all the tested methods to obtain the autocorrelation value, except of "raw ACF", are invariant to the scale of the signal. (b) Sensitivity to the shift of the signal. Same as panel a, except here, the signal is transformed by adding a constant offset, that is, shifted along the *y*-axis. The only two methods invariant to shift transformation are Pearson's correlation, and the ACF value standardized by *z*-scoring. (c) Sensitivity to changes in signal shape. Two versions of a signal based on a repeating rhythmic pattern and a square-wave unitary response are shown on the left. The orange signal was generated using a square wave with a short duty cycle, whereas the purple signal was generated using a long duty cycle (see also Figure 5). Pearson's correlation taken between the original signal and a version time-shifted by the target lag is affected by the shape of the unitary responses, despite their unaltered temporal relative arrangement. The ACF of each signal is shown on the right. The *z*-scored ACF value at the target lag is equal for the two versions of the signal, demonstrating high sensitivity and specificity to periodic recurrence provided by the *z*-scoring standardization.

period can be tagged as "beat-related". This is motivated by the fact that if a response consistently repeats itself with a period P_B, the correlation of the response with a time-shifted version of itself is expected to be maximal when the time shift corresponds to $1P_B$, $2P_B$, $3P_B$, $4P_B$, etc. (see Figure 2).

Next, "beat-unrelated" lags need to be selected, in order to standardize the autocorrelation values obtained from the "beat-related" lags. Here, a similar approach to the magnitude spectrum-based frequency-tagging can be applied, by selecting lags that are not integer multiples of the beat period, yet high autocorrelation values may be

expected at these lags given the design of the stimulus. When the rhythmic stimulus sequence is constructed on an isochronous grid of time points (see Figure S1A), a set of beat-unrelated periods can be chosen, such that they correspond to integer multiples of the grid interval (hence plausible beat periods given the temporal structure of the stimulus).

A set of beat-unrelated lags can be built by taking integer multiples of these periods that do not overlap with the beat-related lags. For example, if the stimulus was constructed on a regular grid with a .2-s interval and one aims to quantify the representation of a beat with period .8 s in the response signal, beat-unrelated lags can be selected as integer multiples of .6, 1.0 and 1.4 s that do not overlap with multiples of .8 s. This example is illustrated using simulated signals in Figure 3 (note that multiples of .4 s were excluded from the set of beat-unrelated lags to match the selection applied to the empirical datasets; see Section 5.3).

As expected, increasing self-similarity of the signal at .8 s leads to gradually increasing mean *z*-scored autocorrelation values at beat-related lags. Critically, the *z*-score remains unchanged when the self-similarity at .8 s is kept constant but the signal is assembled from a unitary response that has a different shape (see Figure 5). These two complementary simulations thus demonstrate both the high sensitivity *and* specificity of this measure to periodic recurrence.

5.2 | Accounting for noise

As illustrated in the previous section, the autocorrelation-based extension of the frequency-tagging approach is valid when applied to simulated signals. However, empirical measurements of physiological signals in general, and scalp recordings of brain activity in particular, inherently contain noise originating from (i) activity unrelated to the processing of the rhythmic stimulus and (ii) artifacts due to, for example sweating, movement, electrode and amplifier noise, etc. (Luck, 2014).

The noise predominant in physiological recordings can be described as broadband, aperiodic, with a characteristic 1/f-like spectrum where power is inversely proportional to frequency (Groppe et al., 2013; He, 2014; He et al., 2010; Miller et al., 2009). This kind of noise can have detrimental effects on the estimated periodic recurrence of a response. Specifically, by systematically distorting the autocorrelation function, 1/f noise can strongly bias the mean *z*-scored autocorrelation observed at beat-related lags. An example is shown in Figure 7a, demonstrating that the more noisy the signal, the more the estimated beat-related *z*-score is "pulled" away from its true value. Importantly, when the noise dominates the signal, the estimated *z*-score converges towards a small non-zero value (Figure 7b).

The detrimental effect of noise can be reduced by using an appropriate noise-correction method. It is important to note that no method can fully "remove" noise. Rather, the effect of noise can be suppressed to an extent given by the quality of the applied method, as well as the nature of the response and noise. Indeed, capitalizing on the prior knowledge of the response and noise is critical to the development of a powerful noise-correction method. For example, as discussed in Section 4.2, beat perception is often investigated using stimuli generated by seamlessly looping a rhythmic pattern. In such cases, the spectrum of the elicited response is expected to contain sharp peaks only at frequencies corresponding to the *pattern repetition rate and harmonics* (see Section 4.3). Since the noise is expected to be broadband, its spectrum should comprise smoothly varying magnitudes across *all* frequency bins (Bach & Meigen, 1999; Norcia et al., 1989; Retter et al., 2021; Strasburger, 1987).

Based on these assumptions, there are several ways to estimate the magnitude spectrum of the noise that can be separated from the response. For example, the magnitude spectrum of the recorded signal can be low-pass filtered, which effectively removes the narrow peaks corresponding to the response, leaving an estimate of the noise spectrum. Likewise, the noise can be estimated by fitting a 1/f-shaped function to the spectrum of the signal (cf., e.g. Donoghue et al., 2020).

The method used in the current paper is based on a recently proposed irregular-resampling auto-spectral analysis (IRASA; Wen & Liu, 2016; Gerster et al., 2022). This method aims to decompose the spectrum of a signal into a periodic component (here corresponding to the response elicited by a periodically repeating rhythmic stimulus) and an aperiodic 1/f component (here corresponding to noise). The estimated magnitude spectrum of the noise can be subtracted from the complex-valued Fourier spectrum of the recorded signal (as illustrated in Figure 8). If the noise estimate has a larger magnitude than the signal at a particular frequency, the complex Fourier coefficients at that frequency can be set to zero. Subsequently, the autocorrelation function can be computed from the noise-subtracted spectrum. As shown in Figure 7c, autocorrelation values obtained after applying the noise-subtraction method are less distorted by noise, compared to values obtained from the raw signal.

The effect of noise can be further reduced by using the a-priori known frequency content of the response. Since all frequencies that do not overlap with the pattern repetition rate and harmonics only capture noise (see Section 4.3), they can be removed from the spectrum by

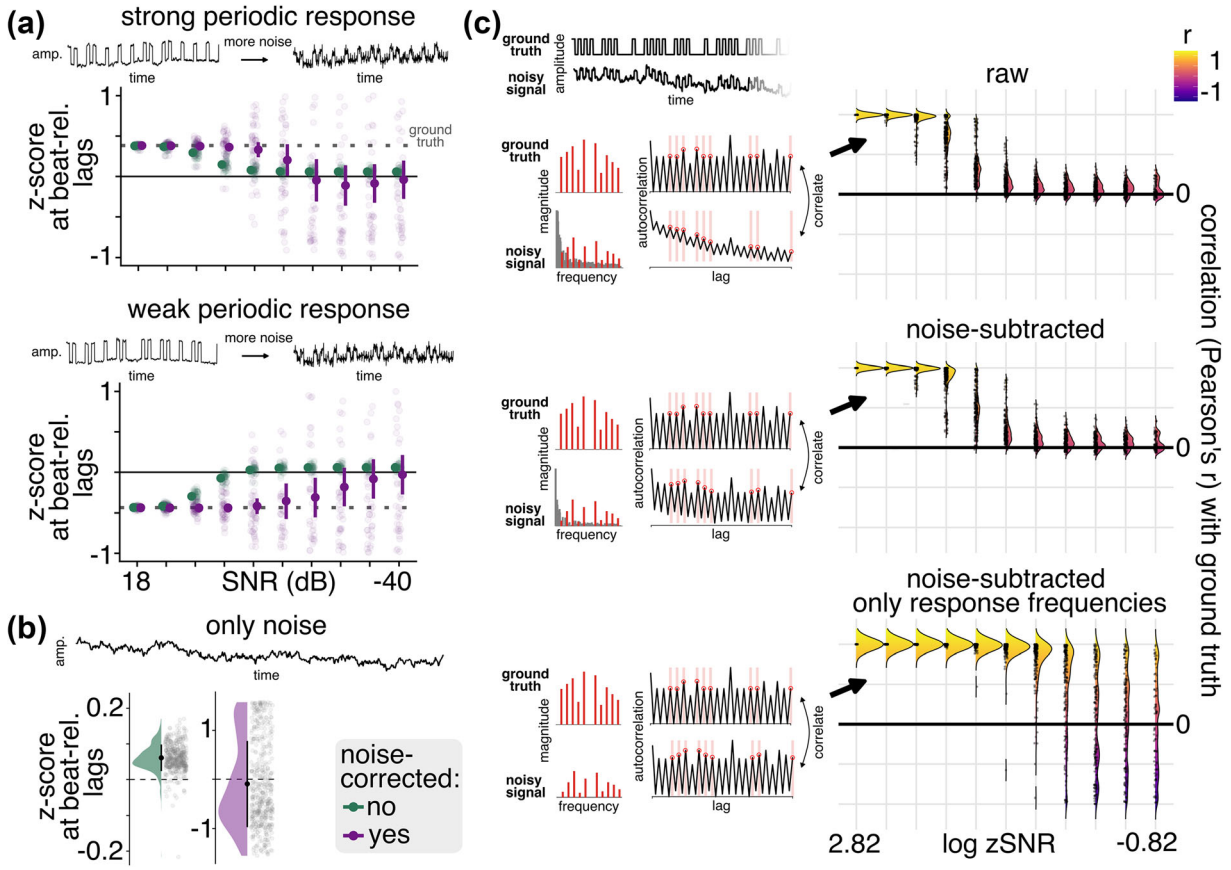


FIGURE 7 Autocorrelation from noisy data. (a) Mean *z*-scored autocorrelation at beat-related lags calculated from a signal generated by summing a simulated response (made up of 20 repetitions of a 2.4-s-long pattern) with different amounts of resting-state EEG noise (see Methods). The autocorrelation was computed either directly from the noisy signal (shown in green) or after noise correction (shown in purple). Individual light circles correspond to *z*-scores obtained from individual simulated signals. Dark circles indicate the mean calculated across simulations separately for each signal-to-noise (SNR) level. Error bars indicate 95% confidence intervals. The horizontal dashed line indicates the ground-truth value calculated from the corresponding response without any noise. The response used in the top panel was chosen to yield a strongly positive *z*-score at beat-related lags, whereas the response in the bottom panel shows a strongly negative *z*-score. Examples corresponding to time-domain segments of the generated signals are shown on the top of each panel. For both panels, the estimated *z*-score remains close to the ground truth value when little noise is present (i.e. the SNR is high). As more noise is added to the signal and the SNR decreases, the *z*-score estimated without noise correction is increasingly biased towards a small non-zero value. When the autocorrelation function is noise-corrected, the estimated *z*-score stays close to the ground-truth value even with lower SNR, and the estimate converges towards zero once the noise dominates the signal, as theoretically expected. (b) Mean *z*-scored autocorrelation at beat-related lags obtained from pure noise (i.e. without any response; example time-domain segment shown on the top). Gray points indicate values obtained from individual simulated signals. The mean and standard deviation across simulations are indicated by the black circle and error bars, respectively. Without noise correction (left plot in green), the observed values are biased towards a small non-zero value (here positive). Applying noise correction removes the bias (although it does increase the variance), and the values converge to zero (right plot in purple). (c) Recovering ground-truth autocorrelation (ACF) values of a simulated response from noisy signals without accounting for noise at all (top), after estimating and subtracting the 1/f-like noise component (middle), and after additionally zeroing-out magnitudes at frequencies that do not capture the response (bottom). A schematic of the time-course, magnitude spectrum and the corresponding ACF of an example signal are shown on the left, separately for each ACF estimation method. The frequencies capturing the simulated response are highlighted in red. The lags where ACF values were extracted are indicated by red vertical lines. The vector of extracted ACF values was correlated with the vector obtained from the ground-truth response without noise. The distribution of the correlation values across all simulated signals is shown on the right, as a function of the observed signal-to-noise ratio (zSNR, see Section 7.2). Individual simulated samples are shown as small black circles. Without noise correction, the correlation quickly drops to zero as the noise level increases. Subtracting the estimated 1/f component before computing the autocorrelation function helps to retain bigger correlation values at lower levels of zSNR. Notably, zeroing-out all frequencies where no response is expected leads to a substantial increase of correlation values even at higher levels of noise.

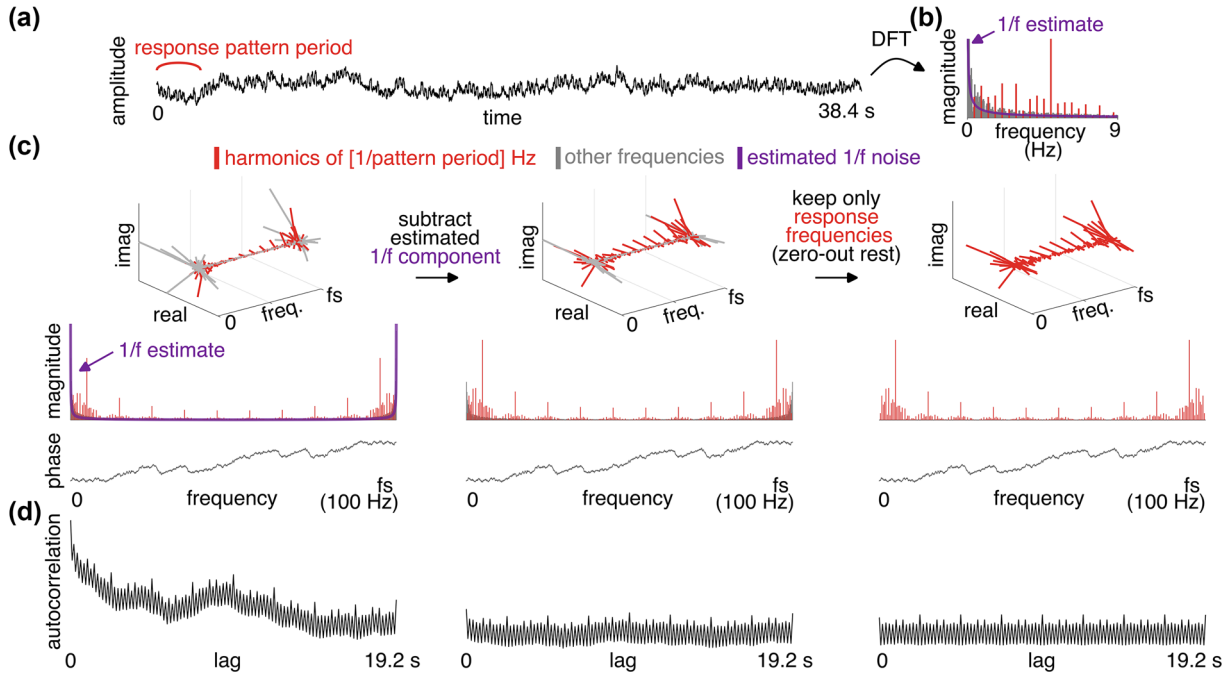


FIGURE 8 Illustration of the noise-correction method. (a) Example noisy time-domain signal comprising a periodic response based on a repeating rhythmic pattern (16 out of 20 repetitions shown for visualization purposes). (b) Magnitude spectrum of the raw signal with response frequencies highlighted in red. The estimated 1/f-like noise component is indicated by a purple line. (c) (Left) Full complex-valued spectrum of the raw signal up to the sampling rate (fs = 100 Hz). The real and imaginary component as a function of frequency is shown on the top. Below, the same information is expressed as magnitude and phase as a function of frequency. As the signal is real, the complex-valued spectrum is conjugate-symmetric around the Nyquist frequency (i.e. half of the sampling rate). The magnitude of the estimated noise component is shown in purple. (Middle) The complex-valued spectrum after the magnitude of the estimated noise has been subtracted, separately for each frequency. This subtraction does not affect the phase. (Right) The complex-valued spectrum after all frequency bins except those capturing the response were set to zero. Again, this operation does not affect the phase spectrum. (d) Autocorrelation functions calculated directly from the complex-valued spectra above.

setting their complex Fourier coefficients to zero (as illustrated in Figure 8). Figure 7c shows that this additional step makes the estimated autocorrelation values robust to even higher levels of noise compared to the noise-subtraction step alone. Similarly, the *z*-score at beat-related lags obtained after applying the two-step noise correction is less biased by noise (Figure 7a), and in cases where the noise dominates the recorded signal, the *z*-score converges towards zero as theoretically expected, rather than a small non-zero value (Figure 7b).

5.3 | Applying autocorrelation to empirical signals

The previous sections have shown how autocorrelation can be used to measure periodic self-similarity of a response to a rhythmic stimulus in a way that is comparable across signals and robust to noise. The validity of the method has been demonstrated using simulated data, which has the benefit of allowing the ground-truth to be known and the properties of the signal to be carefully controlled. What remains to be shown is that the autocorrelation method can be successfully applied to empirical signals. To this end, the method is used here to re-analyze datasets from two previous studies that originally employed frequency-tagging restricted to magnitude spectrum analysis to measure beat-related information in neural activity captured with surface electroencephalography (EEG).

The two studies used a comparable stimulus design, described in detail in the corresponding publications (Lenc et al., 2018, 2022). In short, participants listened to rhythmic sequences created by seamlessly looping a 2.4-s-long rhythmic pattern. The pattern was generated by arranging eight identical pure tones on an isochronous grid of 12 time points separated by .2 s. The rhythm was either made up of low-pitched tones (130 Hz) or high-pitched tones (1236.8 Hz). In one condition, the constituent pattern was considered a "strongly periodic rhythm", since the groups of tones were arranged in a way that closely matches a beat with a rate of 1.25 Hz. This rate was further confirmed as corresponding to the perceived beat by asking the adult participants in the study of Lenc

et al. (2018) to tap the perceived beat while listening to the same rhythmic stimulus in a dedicated session subsequent to EEG recording (with a minority of participants tapping at a rate of 2.5 Hz, i.e. a tempo twice as fast as most other participants).

Neural responses were captured by recording EEG activity from the scalp surface of these adult participants as they were listening to the strongly periodic rhythm in eight 50.4-s-long trials (i.e. the stimulus presented in each trial comprised 21 repetitions of the rhythmic pattern) without any movement. Five 60-s-long trials (i.e. four more pattern repetitions per trial as compared to adults) were obtained for the infant dataset. As expected based on stimulus design, transforming the trial-averaged EEG responses to the frequency domain revealed peaks at frequencies corresponding to the pattern repetition rate (1/2.4 s = .416 Hz) and harmonics (see Figure 9). As discussed in Section 4.3, these peaks effectively isolate time-locked EEG activity that consistently occurred at every repetition of the rhythmic pattern in the stimulus sequence.

Having captured the neural activity related to the processing of the stimulus, the next step is to investigate to what extent the EEG response shows self-similarity at the

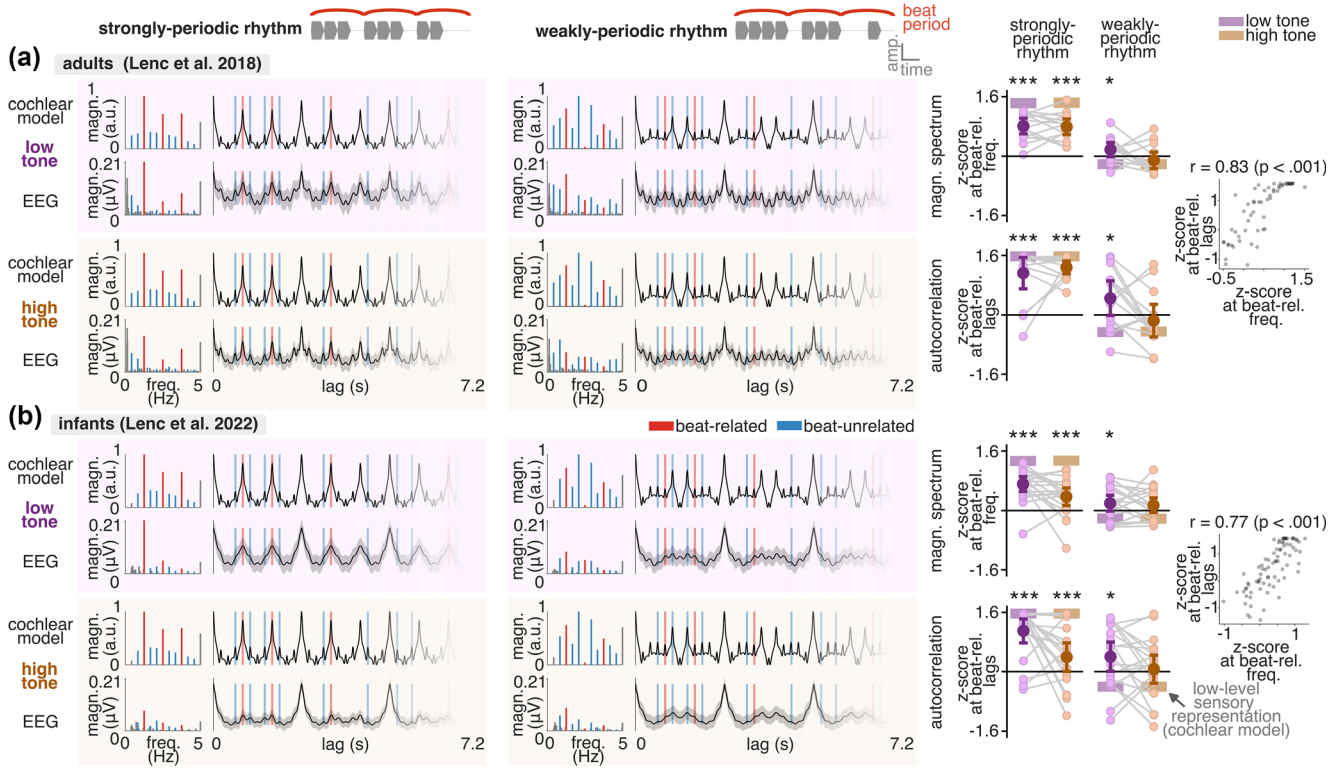


FIGURE 9 Re-analysis of previously published EEG data. (a) Data from healthy adult participants ($N = 14$) (Lenc et al., 2018). (Left) Magnitude spectra and autocorrelation functions of responses to the rhythmic stimuli are shown separately for each rhythm (strongly and weakly periodic) and tone (low and high). For each condition, the top row shows the spectrum and autocorrelation of a lower-level sensory representation obtained from the cochlear model. The bottom row shows the EEG response averaged across participants. Standard error of the mean for the autocorrelation function is indicated using shaded regions. A single repetition of the rhythmic sound pattern used to construct the stimulus sequence in each condition is depicted on the top of the figure, and the corresponding beat period most consistently tapped by participants in response to the rhythm is illustrated by the red curves. Beat-related frequencies/lags are shown in red, and beat-unrelated frequencies/lags used for standardization are depicted in blue. The autocorrelation function was trimmed at 7.2 s for visualization purposes. (Right) Values extracted from EEG responses of individual participants are shown as light-colored individual circles connected by gray lines. Dark-colored circles correspond to the mean response across participants, and error bars indicate 95% confidence intervals (Morey, 2008). The horizontal line segments indicate values extracted from the cochlear model. Asterisks indicate a significant one-tailed *t*-test against zero (FDR corrected, $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$). Mean *z*-scored magnitudes at beat-related frequencies are shown in the top panel, and mean *z*-scored autocorrelation values at beat-related lags in the bottom panel. The scatter plot on the right side of each panel depicts the relationship between beat-related *z*-score values extracted from the EEG responses using the magnitude spectrum (*x*-axis) and autocorrelation (*y*-axis). Each gray circle represents a response from one participant and condition. Pearson's correlation coefficient (r) is shown on the top of the scatter plot. (b) Same structure as panel a, but showing EEG responses from healthy 5- to 6-month-old infants ($N = 20$) (Lenc et al., 2022).

rate of the perceived beat, nested within the cycle of the repeating rhythmic pattern. Given the rate of the most consistently perceived beat, the peaks in the EEG spectrum at the harmonics of 1.25 Hz were tagged as *beat-related*, and the rest of the peaks as *beat-unrelated*.

It should be pointed out that the maximum frequency of interest here was 4.58 Hz. Higher frequencies were excluded to avoid distortion of the EEG spectra by the alpha activity artifact (approximately between 8 and 12 Hz; Luck, 2014; van Diepen et al., 2016; van Diepen & Mazaheri, 2017). The peak at 5 Hz was also discarded from further analysis as it corresponds to the frequency of the time grid on which the rhythmic stimulus was constructed (i.e. shortest inter-onset interval), and is therefore prone to be substantially modulated by features non-specific to periodic recurrence of the signal (see Figure S2). The magnitudes across all selected frequencies were standardized by *z*-scoring as described in Section 4.3 (see Equation 7 in Section 7.2).

Across participants, the beat-related frequencies prominently stood out in the EEG spectra (in contrast to beat-unrelated frequencies), as reflected in their high *z*-scored magnitudes. High positive beat-related *z*-scores indicating prominent periodic recurrence at the rate of the beat were observed in the EEG responses of both adults and infants, irrespective of the pitch of the tone delivering the strongly periodic rhythm (all p-values $<$ .01, one-tailed t-test against 0, see Table S1).

However, as discussed in Section 4.3, these EEG responses may be largely explained by lower level sensory tracking of acoustic features. This was tested by simulating neural responses elicited by the stimulus at the level of the auditory nerve with a well-established biomimetic model of the auditory periphery (Slaney, 1998). The output of this "cochlear model" was transformed into the frequency domain, revealing positive *z*-scored magnitude at beat-related frequencies, which was never exceeded by the corresponding EEG responses (all p-values $>$.05 across datasets and tone conditions, one-tailed t-test, see Table S1). Firstly, this shows that for the strongly periodic rhythm, the perceived beat matched the most prominent periodic recurrence already present in the physical structure of the stimulus. More importantly, this case highlights how the neural representation of the beat can be confounded by lower level tracking of acoustic features, in the case of strongly periodic sensory inputs.

To control for this acoustic confound, the stimulus sequence in the second condition was made of a repeating "weakly periodic rhythm", whereby the same number of tones was positioned based on the same periodic grid of time intervals as for the strongly periodic rhythm, but in a way that did not systematically match any plausible periodic beat. Despite this, the tapping responses confirmed that participants consistently perceived a beat at the same rate as for the strongly periodic rhythm (this result has been corroborated by several other studies, e.g. Nozaradan et al., 2012, 2018).

The fact that this induced beat did not correspond to an acoustically prominent periodicity in the weakly periodic stimulus was confirmed by the low *z*-scored magnitude of beat-related frequencies in the spectrum of the lower level sensory response simulated with the cochlear model. Yet, when the rhythm was delivered with low-pitched tones, the EEG responses showed an enhancement of beat-related frequencies, indicated by significantly positive *z*-scores above the values obtained from the cochlear model. This effect was consistently observed in the responses from both adults ($t_{13} = 2$, $p = .04$) and infants ($t_{19} = 1.94$, $p = .04$, one-tailed t-test against 0) and did not occur in the high-tone condition (all p-values $>$.05, see Table S1).

These results indicate that low-pitched rhythmic sounds may engage internal transformation processes that selectively emphasize the beat in the neural representation when it is not prominent in the sensory input. In other words, when stimulated with "bass-like" sounds, the neural system seems to "periodize" the internal representation of the stimulus towards the perceived beats even in the absence of overt body movement.

However, these results could have been driven by response properties that are unspecific to periodicity, thus constituting a false positive. Specifically, an equivalent unitary brain response could have been uniformly elicited by every tone of the weakly periodic rhythm *within conditions*, yet with differences in its shape *across conditions* (and with respect to the cochlear model), thus giving rise to differences in beat-related *z*-scores across conditions as observed here (see Figure 5 for a simulated example).

A powerful way to test whether the magnitude spectrum–based results reported above constitute a false positive is to capitalize on the new autocorrelation-based extension of frequency-tagging developed in the current paper. Namely, autocorrelation offers a way to measure periodic self-similarity invariant to the particular shape of the recurring signal. Therefore, confirming the results obtained with the magnitude spectrum–based analysis using the autocorrelation-based analysis would provide critical confirmation for a true selectively enhanced representation of periodicities at the rate where human adults tend to tap the beat when listening to the weakly periodic rhythm, particularly when the rhythm is delivered by low-pitched sounds.

Based on the rate of the most consistently perceived beat, lags corresponding to integer multiples of

1/1.25 Hz = .8 s were tagged as beat-related lags. Beat-unrelated lags were selected as .6, 1.0 and 1.4 s and their integer multiples, which all in turn correspond to integer multiples of the periodic grid of time intervals used to construct the rhythmic stimulus (see Section 5.1). Lags overlapping between the two sets (beat-related lags and beat-unrelated lags) were excluded from the analysis. In addition, since some participants tapped the beat rate of 2.5 Hz, the beat-unrelated lags overlapping with integer multiples of 1/2.5 Hz = .4 s were excluded as well. This way, the mean *z*-score at beat-related frequencies was optimized to measure periodic recurrence at the beat rate most commonly tapped by participants, yet still accounting for another, perhaps less perceptually salient metric layer (note that the results remained unchanged even when beat-unrelated lags overlapping with multiples of .4 s were not excluded).

As shown in Figure 9, for the cochlear model output, the mean *z*-scored autocorrelation at beat-related lags was highly positive for the strongly periodic rhythm, and below zero for the weakly periodic rhythm. In other words, autocorrelation converged with the magnitude spectrum analysis at showing that the acoustic structure of the strongly periodic rhythm was characterized by prominent periodic recurrence at the rate of the perceived beat, which was not the case for the weakly periodic rhythm.

For the EEG responses, the autocorrelation function was obtained after applying the noise-correction procedure as described in Section 5.2. First, the 1/f-like noise component was estimated using IRASA, separately for each condition and participant, and subtracted from the complex-valued spectrum of the response. Then, magnitudes at all frequencies, except harmonics of the pattern repetition rate (1/2.4 s = .416 Hz), were set to zero. The autocorrelation function calculated from the noise-corrected spectrum was used to extract values at beat-related and -unrelated lags. For the strongly periodic rhythm, beat-related *z*-scores in the EEG were fully explained by lower level sensory tracking of acoustic features, as captured by the cochlear model (all *p*-values >.05 across datasets and tone conditions, one-tailed *t*-test, see Table S1). On the other hand, the EEG responses to the weakly periodic rhythm delivered by low-pitched tones revealed significantly enhanced beat-related *z*-scores in the data obtained from both adults ($t_{13} = 2.10$, $p = .04$) and infants ($t_{19} = 2.14$, $p = .03$, one-tailed *t*-test against 0). Such significant enhancements of the beat periodicity were not observed for the high-pitched tones (all *p*-values >.05, see Table S1).

Thus, for the two re-analyzed datasets, the magnitude spectrum- and autocorrelation-based analyses provided convergent results. Importantly, the *z*-scores at beat-related frequencies (magnitude spectrum-based analysis) and *z*-scores at beat-related lags (autocorrelation-based analysis) were strongly correlated across participants in the data from adults ($r = .83$, $p < .001$) as well as infants ($r = .77$, $p < .001$) (see Figure 9). Together, these results demonstrate two important points. First, the autocorrelation-based analysis can be successfully applied to empirical noisy data such as surface EEG (even recorded in human infants), while having sufficient sensitivity to reveal differences between conditions. Second, the convergence between the two analyses provides strong evidence that the results obtained by analyzing these particular datasets using magnitude spectra cannot be simply explained by properties of the EEG responses that are non-specific to periodicity. For further discussion of the theoretical implications of these results, we encourage readers to refer to the original publications (Lenc et al., 2018, 2022).

To show that the autocorrelation approach can be applied to a range of time-varying responses beyond EEG, the same analysis was performed on the finger-tapping data obtained from adult participants during a session following the EEG recording. These responses corresponded to time-series of mechanical vibrations elicited by the tapping finger on the underlying surface of a tapping box. As shown in Figure S3, the observed *z*-scores at beat-related frequencies and beat-related lags were again strongly correlated across participants, conditions, and trials (Pearson's $r = .93$, $p < .001$). This result further corroborates the convergence between the magnitude spectrum–based and the autocorrelation-based analysis, as indicated already through the analysis of the EEG data above.

6 | DISCUSSION

6.1 | Measuring beat-related information in empirical signals as a second-order isomorphism

Rhythm perception often relies on mapping an external sensory input onto an internal temporal reference consisting in a periodic beat (Honing & Bouwer, 2018; Large, 2008; London, 2012). The nature of this mapping can be investigated by measuring how the periodicity of the perceived beat is represented in a rhythmic sensory input vs. behavioral/neural response to this input. However, this periodicity can take many forms in empirical signals, as diverse as slow ramping versus transient burst of neuronal firing rates produced at fixed positions within each cycle of the beat period (see e.g. Gámez et al., 2019; Merchant et al., 2014).

Moreover, the beat has been often conceptualized as a series of regularly recurring, homogeneous psychological events (Honing & Bouwer, 2018; Large & Snyder, 2009). But it remains to be determined whether these events correspond to discrete points in time (Cohn, 2020), or whether they may have some substantial temporal span, thus better characterized as beat bins (Danielsen et al., 2023; Large & Palmer, 2002). An equally plausible way to describe the beat could be as a continuous phase representing progress through each cycle of the beat period (Cannon, 2021).

Crucially, while the exact nature of the beat is not yet definitively resolved, all these perspectives sustain the basic notion of beat as periodic recurrence. Therefore, the current paper argues that conceptualizing beat as a *periodic self-similarity* (i.e. *periodic recurrence*) constitutes a promising framework to measure how beat-related information is contained in empirical signals while minimizing underlying assumptions (see also Lenc et al., 2021). Such a parsimonious approach allows to study the mapping between external sensory inputs and perceived beat from a functional perspective, without additional assumptions about the particular shape of the recurring trajectory.

Based on this conceptual framework, the current paper proposes a new methodological approach whereby the extent to which neural/behavioral responses contain beat-related information is measured as a *second-order isomorphism* (Kriegeskorte et al., 2008; Shepard & Chipman, 1970). More specifically, the methods discussed here assess the extent to which the response repeats itself at a rate compatible with the hypothesized perceived beat rate.

6.2 | Highlighting strengths and going beyond limitations of frequency-tagging based on magnitude spectrum analysis

Among the variety of methods that have been proposed to capture periodic recurrence (see Lenc et al., 2021 for a review), frequency-tagging has been increasingly adopted in the neuroscience community over the last decade (Bouvet et al., 2020; Celma-Miralles et al., 2016; Celma-Miralles & Toro, 2019; CSifuentes-Ortega et al., 2022; Cirelli et al., 2016; Lenc et al., 2018, 2020, 2022; Li et al., 2019; Nozaradan, Mouraux, et al., 2016; Nozaradan, Peretz, & Keller, 2016; Nozaradan et al., 2011, 2012; Nozaradan et al., 2018; Okawa et al., 2017; Sifuentes-Ortega et al., 2022; Tal et al., 2017; Tierney & Kraus, 2014).

Our simulations corroborate the high *sensitivity* of frequency-tagging to periodic recurrence as measured from the magnitude spectrum of slowly fluctuating, physiologically plausible responses, without the need to rely on additional assumptions about the particular shape of the recurring trajectory. Following up on previous work (Nozaradan, Keller, et al., 2017; Lenc et al., 2021), we re-iterate the critical role of taking stimulus features into account in order to prevent lower level sensory confounds and show how this can be effectively accomplished by measuring periodic recurrence in the stimulus with magnitude spectrum-based analysis.

Importantly, we also discuss the caveats of the method, particularly related to the limited *specificity* with respect to periodic recurrence as measured using the magnitude spectrum of the signal (Rajendran et al., 2017; Rajendran & Schnupp, 2019). To address this well-known limitation, alternative approaches had already been put forward, consisting for instance in decomposing the signal into periodic but non-sinusoidal components that are not specified a-priori but estimated directly from the signal (Leman & Naveda, 2010; Sethares & Staley, 2001).

Here, this limitation is addressed by introducing a novel analysis based on autocorrelation, generally inspired by recent work on neural processing of musical rhythm (Herff et al., 2020). Our method offers a straightforward and deterministic way to quantify periodic recurrence in a signal at a given rate of interest, without the need to choose between various algorithms that search for periodic components (c.f., e.g. Sethares & Staley, 2001).

Using simulations, we demonstrate that the autocorrelation-based analysis is both *sensitive* and *specific* to periodic recurrence, showing invariance to changes in the shape of the signal that preserve periodic self-similarity. Critically, we show how the method can be used to directly compare beat-related information in the physical features of a sensory input and in the associated response.

More generally, the capacity to quantify and compare periodic recurrence across a range of signals provides a pivotal tool for studying the gradual emergence of beat representation across different brain areas and processing stages. For example, the autocorrelation-based analysis could be used to characterize the progressive transformation from a representation of lower level sensory features in the peripheral receptors to an abstracted representation in high-level cortical regions that would predominantly reflect a periodic beat, which would itself be exclusively represented in the overt body movement of the participant (Nozaradan et al., 2018, Nozaradan, Mouraux, et al., 2016, Rajendran et al., 2017, 2020). In fact, as long as they have sufficient temporal resolution, the compared signals can come from very different empirical modalities, such as time-varying acoustic

envelope (Ding et al., 2017), output of a computational model (Zuk et al., 2018), spiking rate (Chang, 2015; Merchant & Averbeck, 2017; Rajendran et al., 2017), oscillatory power (Fujioka et al., 2015; Herff et al., 2020; Iversen et al., 2009; Merchant & Bartolo, 2018), large-scale field potentials (Nozaradan, 2014; Nozaradan, Mouraux, et al., 2016; Tal et al., 2017) or movement velocity (Patel et al., 2009; Schachner et al., 2009).

Overall, the autocorrelation-based analysis thus appears promising to shed light on beat perception by offering a functional perspective that can connect key elements of systems neuroscience, i.e., the external input, the associated brain activity, the elicited behavior as well as computational models (Kriegeskorte et al., 2008; Kriegeskorte & Wei, 2021).

6.3 | What stimulus features to take into account?

While we highlighted the issue of lower level sensory confounds, and the critical role of taking the stimulus into account, we did not specify *which* exact features of the stimulus should be compared to the neural/behavioral responses. This depends on the particular experimental context and the nature of the stimulus. In principle, the neural system may show a representation of any acoustic feature. In other words, the neural activity may be a function of that feature (Brodbeck et al., 2018; Daube et al., 2019; Hamilton et al., 2021; Nourski & Howard, 2015; Wang, 2018). If the feature varies with high periodic recurrence over time, the associated response will, likewise, show high periodic recurrence, thus conflating lower level sensory processing with internal beat representation (Nozaradan, Keller, et al., 2017).

Determining the relevant acoustic features may not be trivial, especially with more ecologically valid stimuli. For tightly controlled stimuli, acoustic amplitude envelope or frequency may be the only feature modulated at a time scale relevant for beat perception (e.g. Lenc et al., 2018; Nozaradan et al., 2018). However, for more complex stimuli, one may choose to control for variations in timbre, pitch and perhaps even higher-level attributes such as harmony and different combinations thereof (Hannon et al., 2004; Keller & Schubert, 2011; Povel & Okkerman, 1981; Toiviainen & Snyder, 2003).

Selection and extraction of the relevant sensory features remain an empirical endeavor, which can be facilitated by using computational models with various degrees of physiological plausibility (Daube et al., 2019; Weineck et al., 2022). Relatedly, the choice of a particular computational model must be justified by the research question being addressed. For example, when re-analyzing the empirical datasets in the current paper, we used a biomimetic model which simulates sound processing up to the level of the auditory nerve (Slaney, 1998). By comparing the surface EEG activity to this cochlear model, we could evaluate the selective enhancement of the beat-related periodicity beyond *what is already present* in the auditory nerve. However, it is important to note that transformations of sound-evoked activity *before* it reaches the auditory nerve are not in any sense less relevant. For example, nonlinear properties of the cochlea may produce rudiments of periodic recurrence as compared to the acoustic input itself (Rajendran et al., 2020; Wojtczak et al., 2017; Zuk et al., 2018).

In sum, considering different representations and transformations of the sensory input and comparing their periodic recurrence can isolate processes which contribute to the emergence of the beat representation throughout the nervous system (Hamilton et al., 2021; Mattioni et al., 2022; McDermott & Simoncelli, 2011; Sankaran et al., 2018).

6.4 | The frequency-tagging approach as implemented through autocorrelation-based and magnitude spectrum-based analyses

In the current paper, we also directly examined how noise impacts the measures obtained with the autocorrelation-based analysis, and we provided concrete ways to mitigate these detrimental effects. Explicitly characterizing the potential biases and confounds due to noise and SNR are critical for any methodological approach that may be applied to physiological signals (e.g. van Diepen & Mazaheri, 2018). Yet, such endeavors remain scarce, particularly in the field of timing and music neuroscience.

Our simulations indicate that the autocorrelation-based analysis is robust to noise, particularly when the impact of noise is suppressed by capitalizing on the same principles exploited in magnitude spectrum-based implementations of frequency-tagging to achieve high segregation between signal and noise (Norcia et al., 2015; Rossion et al., 2020). In particular, tightly controlling the temporal structure of the input allows for objectively isolating the response concentrated in a small set of a priori defined frequency bins (see Section 4.2). For this reason, the autocorrelation approach can be thought of as an extension of frequency-tagging. Indeed, the DFT and autocorrelation are linked by an invertible transformation, essentially reflecting the same information albeit in different, potentially complementary, formats.

6.5 | Complementarity of autocorrelation-based and magnitude spectrum–based analyses to disentangle multiple beat layers: Investigating musical meter

The different ways in which magnitude spectrum and autocorrelation reflect similar information offer robust complementarity across the two methods, particularly in the context of studying *meter* perception. In the current paper, we mainly considered the internal representation of a single periodic layer (the beat) for simplicity.

However, as mentioned in Section 1, listening to music often induces perception of a meter, which consists of multiple nested layers of periodic recurrence (beat layers) forming a coherent temporal structure (Cohn, 2020; London, 2012). For example, the tapping data of Lenc et al. (2018) indicate that a meter with at least two periodic beat layers may have been perceived by the participants, who were then constrained by the behavioral task to select and tap one of these layers (see Section 5.3) (for similar observations, see, e.g. McKinney & Moelants, 2006; Martens, 2011). In fact, it has been argued that perception of a single periodic layer representing the main beat ("the" beat) may be an idealization that lacks unequivocal support from empirical data (Brochard et al., 2003; Cohn, 2014; Large et al., 2002; Repp, 2008; Repp & Jendoubi, 2009; Toiviainen & Carlson, 2022).

The magnitude spectrum and autocorrelation implementations of frequency-tagging both inherently capture periodic recurrence at multiple rates simultaneously, thus allowing the representation of a particular meter as a whole to be quantified in the analyzed response. On the one hand, tagging *harmonic frequencies* of, for instance, 1.25 Hz (e.g. 1.25 Hz ×1, ×2, ×3, ×4, etc.) within the magnitude spectrum includes periodic recurrence not only at .8 s (i.e. 1/1.25 Hz) but also at *faster* rates corresponding to integer divisors of .8 s (i.e. the metrical layers that *could be nested* within this period), i.e. at .4 s (.8 s/2), .2 s (.8 s/4), etc. This is because all harmonics of these faster layers coincide themselves with the harmonics of the slowest layer. On the other hand, tagging integer *multiples of a lag*, for example .8 s (i.e. .8 s ×1, ×2, ×3, ×4, etc.) in the autocorrelation function, captures periodic recurrence at *slower* rates corresponding to metrical layers that *could nest* .8 s within their period, for example at 1.6 s (.8 s ×2), 2.4 s (.8 s ×3), etc.

While both methods allow the internal representation of a meter to be captured without unnecessary assumptions about the relative status of its constituent nested periodic layers, determining which particular layer is predominantly reflected in the analyzed response poses a challenge. Nevertheless, combining the two methods with specific sub-selections of frequencies/lags allows *partially* dissociating between different nested metrical layers that may be driving the experimental effect. This complementarity lies in the fact that autocorrelation contains a subset of lags that *only* pick up self-similarity at rates corresponding to faster metrical layers without being influenced by the slower metrical layers. Conversely, magnitude spectrum contains a set of frequencies that are *only* sensitive to periodic self-similarity at the rates of slower but not faster metrical layers (as further illustrated in Figure S6).

Hence, the analyses based on magnitude spectrum and autocorrelation are expected to provide convergent results if the selection of frequencies and lags captures the metrical layer(s) that are prominently represented in the analyzed response, as exemplified in the strong correlation between the two methods we observed in our data.

6.6 | Corroborating conclusions of previous studies

Our analysis corroborates the results obtained with magnitude spectrum–based analysis in the original papers (Lenc et al., 2018, 2022), providing evidence that these original effects were not merely driven by changes in signal shape that were not specific to periodic recurrence (Rajendran et al., 2017; Rajendran & Schnupp, 2019). In saying that, one difference from the original papers was the lack of significant enhancement of beat periodicities in the EEG responses to the weakly periodic rhythm delivered by high tones. In the case of the magnitude spectrum-based analysis, this may be partly explained by a slightly different selection of frequencies (excluding the grid rate frequency; see Figure S2). More importantly, this may be due to a more stringent test of the beat-related z-score against zero rather than the beat z-score from the cochlear model. While directly comparing the z-score of a response to some baseline (here cochlear model) is valid in theoretical noise-free scenarios, caution is required when analyzing noisy physiological signals. In particular, our simulations reveal that in situations with a low SNR, the noise-corrected z-score at beat-related lags is biased towards zero, and the same can be argued for z-scores at beat-related frequencies. Consequently, in cases where the baseline against which the response is compared has a negative z-score value, noise may spuriously enhance the z-score value of the response by making it less negative.

However, once aware of the detrimental effects of noise, one can adjust the statistical analyses accordingly to make sure the results are robust. For example, when multiple experimental conditions are compared, one

possibility is to empirically estimate the SNR from the data, and use it as a regressor in the statistical model. However, in scenarios with an insufficient amount of data, a reasonable alternative is to make sure that the significant contrasts cannot be simply reproduced when taking the SNR as a dependent variable instead of the beat-related *z*-score (for an example, see Lenc et al., 2018, 2022). These suggestions thus point toward general recommendations for optimization of stimuli and data acquisition, to maximize SNR of the collected data (Boudewyn et al., 2018; Luck, 2014).

6.7 | Limitations and cautionary notes

The simulations conducted in the current paper were designed to yield insights directly relevant to the analysis of real-world empirical data, while maintaining control over the relevant parameters. For instance, rather than artificially generating noise (e.g. van Diepen & Mazaheri, 2018), we opted for sampling it from an actual EEG recording. Likewise, we demonstrated the validity of the method using a unitary response with a physiologically plausible shape.

However, the shape of the unitary response was not estimated from real data, and, in fact, the mere concept of unitary response may be questionable in the context of highly nonlinear neural systems (Ahrens et al., 2008; David et al., 2009; Doelling et al., 2019; Drennan & Lalor, 2019). While this does not challenge the validity of our findings, it is important to acknowledge that the data simulated in the current study may be different from real-world data acquired in various experimental situations. Therefore, the implementation of the autocorrelation approach presented here should be taken as a proof of concept, rather than as a "recipe" for a "plug-and-play" analysis pipeline.

Instead, we provide a set of plausible and validated analytical tools, while pointing out their strengths and limitations, thus opening to further development depending on the requirements of the particular application. For example, while we show how the SNR can be empirically estimated from data, and we use simulations to demonstrate that it affects the autocorrelation method in a predictable way, we do not offer a minimum value of SNR that should be achieved in an empirical experiment to guarantee interpretable data. This is because the absolute SNR may depend on many variables including the number of harmonics taken into account, as well as the distribution of response magnitude across these harmonics. Indeed, different kinds of responses may predominantly project to different regions of the spectrum. For instance, transient burst-like periodic responses will contain more energy concentrated at higher harmonics, whereas slowly varying periodic responses will mainly contain prominent peaks at lower harmonics (Retter et al., 2021; Zhou et al., 2016). Likewise, signals from different empirical modalities may be susceptible to certain idiosyncratic artifacts that distort the spectrum of the response in a particular way (e.g. alpha activity in EEG). All this should be taken into account to fine-tune the selection of frequencies and lags for a particular application of the frequency-tagging as implemented through magnitude spectrum or autocorrelation analyses.

Similarly, we demonstrated the validity of the noise-correction procedure using a particular tool (IRASA) to estimate the noise spectrum from data. However, there are several available approaches to estimate 1/f-like noise component from signals, with different strengths and weaknesses which are beyond the scope of this paper (Gerster et al., 2022). Again, the decision to use a particular method to capture the noise component should be informed by the nature of the data and noise in the particular application.

Yet, it should be noted that there are practical recommendations regarding the duration of the analyzed signal, particularly pertinent to the noise-correction procedure described in Section 5.2. Specifically, to minimize the impact of noise, one needs to (i) reliably estimate the 1/f-like noise component for subtraction and (ii) ensure that the response is concentrated into narrow frequency bins that are clearly separable from frequency bins that are driven solely by noise, which can be subsequently set to zero (see Sections 5.2 and 7.2 for further details of this two-step procedure). This separation depends on the spectral resolution, which is determined by the duration of the analyzed signal (and hence by the number of repetitions of the rhythmic pattern therein).

For example, when a rhythmic pattern is presented only once, with no seamless repetition, the corresponding spectrum contains no additional frequency bins between the frequencies of interest (i.e. the inverse of pattern duration and harmonics). In such a case, the magnitudes captured at the frequencies of interest would receive equal contribution from the response and the noise (Bach & Meigen, 1999). Two seamless repetitions of the rhythmic pattern would yield one frequency bin separating the neighbouring frequencies of interest; three repetitions would yield two bins in between the frequencies of interest, etc. In other words, the higher the number of repetitions, the greater the ability to separate signal from noise. However, in practice, a higher number of repetitions comes at the expense of participant compliance and vigilance due to increased trial duration. Therefore, in most experimental designs, SNR can be bolstered by finding a reasonable compromise between (i) the number of

seamless rhythmic pattern repetitions within a trial, (ii) the quality of the recorded data and (iii) the number of trial repetitions averaged per condition (Norcia et al., 2015).

Finally, it is important to keep in mind that the minimal required number of seamless pattern repetitions should be also determined by the functional phenomenon of interest. For example, beat representation may take several repetitions of a rhythmic pattern to build up (Su & Pöppel, 2012; Tal et al., 2017), and this time course is thus relevant to be accounted for when designing an experiment.

Besides validating the new approach on simulated data where ground-truth is known, we also demonstrated its plausibility with empirical data obtained in two previous studies. Despite the relatively small number of participants ($N = 14$ adults and $N = 20$ infants), convergent results were observed across these two independent datasets. While the statistical power of the frequency-tagging approach remains to be systematically evaluated in future work, the current observation of significant effects consistent with the effects originally reported in each respective study provides support to the claim that the autocorrelation-based analysis is sufficiently sensitive and robust to be applied to limited and inherently noisy physiological data.

6.8 | Neural and behavioral responses as complementary ways to study perception

Probing the internal representation of a high-level perceptual phenomenon such as the beat is a nontrivial endeavor. In line with researchers from other fields (Kappenman & Luck, 2011; Moshel et al., 2022; Rossion et al., 2020; Rossion et al., 2023), we argue that relying solely on overt behavioral responses may give an incomplete image of the internal representations due to the intrinsic constraints and number of internal processes that may contribute to the behavioral response eventually executed by the system.

In contrast, measures of brain activity can circumvent several of these constraints, thus isolating aspects of the processes related to the internal representation of interest which otherwise could be masked or distorted. However, it is important to note that neural responses (even at the level of a single neuron) are also likely to be driven by a range of processes (Grootswagers et al., 2019; Mattioni et al., 2020; Merchant et al., 2014; Norman-Haignere et al., 2015). In this view, both behavior *and* neural activity should be considered as incomplete, yet complementary windows into the underlying system, each with its own set of advantages and drawbacks (Krakauer et al., 2017; Niv, 2021; Rossion et al., 2020). Analyzing a range of behavioral and neural responses acquired with a variety of tools and in many contexts may therefore prove an optimal way to learn about the nature of beat perception.

While the methods described here allow researchers to go a step forward by capturing beat-related information in neural activity, it is important to acknowledge that our methods are not fully dissociated from behavior. In fact, they are informed by behavioral measures in terms of identifying a particular rate for the perceived beat as elicited in response to a particular external input in a particular context. This is similar to most methodological approaches that study the neural underpinnings of high-level perception (Kriegeskorte et al., 2008; Norman-Haignere et al., 2015; Rossion et al., 2020). For example, capturing the internal representation of a visual category from brain activity relies on the analysis of neural responses to a set of images that either contain the visual category or not (Grootswagers et al., 2019; Hagen et al., 2020). This stimulus categorization is often based on behavioral responses of the researchers themselves when they build the set of stimulus images, but could be equally well determined by relying on consistent rating responses of a pool of participants (Grootswagers et al., 2022; Norman-Haignere et al., 2015; Shatek et al., 2022; Wardle et al., 2020), which is similar to the approach we propose in the current paper.

6.9 | Beyond isochronous grid-based and repeating rhythms

The current work mainly focuses on situations where the perceived beat is induced by rhythmic patterns that are periodically repeated to form long sequences, and whose structure is based on an isochronous temporal grid. While such stimuli constitute an optimal setting to capitalize on the advantages of the frequency-tagging approach in its magnitude spectrum and autocorrelation implementations (such as robustness to noise), it is worth pointing out that the methods can be generalized to other scenarios.

For instance, even if the stimulus consists of a sequence made of non-repeating rhythmic patterns, beat-related frequencies and lags can be selected in the same way as for stimuli made of looped rhythmic patterns, as long as it can be assumed that the rate of the internal beat is *fixed* at an *exact known value* throughout the whole duration of the analyzed response. Examples of such stimuli comprise, for example, artificially generated non-repeating sound sequences constructed on an isochronous time grid (see, e.g. Lenc et al., 2020), or quantized music without tempo fluctuations.

However, using non-repeating rhythmic stimuli makes the selection of beat-unrelated lags/frequencies less straightforward. If the stimulus is non-repeating but nevertheless constructed on an isochronous time grid, the beat-unrelated lags can still be selected in the way described here, since the autocorrelation peaks would be expected at integer multiples of the grid interval. However, as illustrated in Figure S7, tagging beat-unrelated frequencies may turn challenging because the precise location of peaks in the spectrum driven by the sensory input is not apparent. Careful analysis of peaks emerging in the magnitude spectrum of the sensory input as well as the elicited response is required to ensure that an appropriate set of beat-unrelated frequencies is selected to obtain a valid estimate of periodic recurrence.

What about situations where the rate of the internal beat may slowly fluctuate in a known/controlled manner? For instance, a rhythmic sequence that induces internal beat representation as captured by behavioral motor responses may gradually change in tempo over time. Indeed, humans readily adjust the rate of the internal beat to maintain tight temporal coordination with a tempo-changing external stimulus (Van Der Steen et al., 2015). The two implementations of the frequency-tagging approach presented in the current paper have not been specifically designed to capture the representation of a tempo-changing beat, since they specifically rely on signals' stationarity (especially stationarity of the perceived beat throughout presentation of the rhythmic stimulus).

Nevertheless, it may be possible to capitalize on recent advances in the development of time-warping methods (Chemin et al., 2018; Merchant et al., 2015; Perez et al., 2013), which could be used as a pre-processing step to rescale the temporal structure of the signal in a way that restores stationarity. However, while an in-depth discussion of the assumptions and potential pitfalls of different time-warping methods is beyond the scope of the current paper, it is important to be aware of any caveats related to the chosen time-warping method when interpreting results.

Finally, the current work only deals with capturing internal representations of isochronous beats, that is, beats comprising recurring evenly spaced time intervals. While perception of isochronous beats is pervasive in musical behaviors around the world (Mehr et al., 2019; Savage et al., 2015), increasing evidence suggests that perception of beat and meter formed of uneven time intervals is far from peculiar over the globe (Polak & London, 2014; Polak et al., 2018), highlighting the remarkable ability of the human brain to map complex internal representations (Hannon & Trehub, 2005; Holt & Lotto, 2010; Ley et al., 2014). Here again, the methods as presented in the current paper have not been specifically designed to capture representations of non-isochronous beats.

However, this methodological constraint in no way downplays the significance of the work aimed at shedding light on the perception of non-isochronous beats (Hannon & Trainor, 2007; Jacoby et al., 2024; Jacoby & McDermott, 2017; Polak & London, 2014). On the contrary, taking advantage of all suitable methods to study the rich musical behaviors across cultures and traditions is critical to move a step forward from models in music sciences that have been built on long-standing Western biases, towards a more comprehensive and accurate understanding of human perception (Jacoby et al., 2024; Sauvé et al., 2023).

6.10 | Magnitude spectrum–based and autocorrelation-based analyses do not ignore phase

It has been argued that the magnitude spectrum–based implementation of the frequency-tagging approach ignores temporal properties of the response encoded in its phase spectrum that may be critical to quantifying how much the response reflects a periodic beat (Cameron et al., 2019; Rajendran & Schnupp, 2019; Rosso et al., 2021, 2023; Tal et al., 2017). Counter-arguments to this claim have been already discussed elsewhere (Lenc et al., 2019) (but see also relevant papers addressing this issue in the wider frequency-tagging community, e.g. Bach & Meigen, 1999; Norcia et al., 2015; Rossion et al., 2020). However, it seems relevant to re-iterate here several key points that may remain under-acknowledged in the field.

First and foremost, the spectrum of a periodically repeating response only contains peaks at harmonics of the response repetition rate. The relative distribution of magnitudes and phases across these harmonics is determined by the particular shape of the repeating response. Critically, as discussed in the current paper, the periodic repetition itself rather than the shape of the recurring response is relevant to measure the prominence of periodicity corresponding to the beat perceived in time-varying signals. It follows from the above that measuring beat-related information (namely prominence of beat-related periodicity) in a signal can be achieved irrespective of the relative distribution of magnitudes and phases across the beat-related harmonics, as demonstrated in Section 4 (see also Figures 3 and 4).

Second, the claim that the magnitude spectrum–based implementation of frequency-tagging ignores temporal information may stem from the false assumption that the method works with a spectrum obtained in a similar way

as it is in time-frequency decompositions typically used to analyze oscillatory responses in neuroscience (Keil et al., 2022). This is incorrect, since frequency-tagging does not compute and average the magnitude spectrum of successive short time windows, which would indeed discard phase information. Rather, the spectrum is computed from a *long signal* that captures many repetitions of the beat period (Bach & Meigen, 1999). This means that any temporal jitter or phase inconsistencies in the perceived beat, which can be, in fact, essentially understood as reduced periodic self-similarity, would be reflected in lower magnitude at beat-related frequencies (due to spectral leakage). A reduction of peaks at beat-related lags due to reduced phase stability can be likewise observed in the autocorrelation function, as demonstrated in Figure 2 of the current paper, and further corroborated by Figure S5.

Highlighting the sensitivity of frequency-tagging to phase-stability in both the magnitude spectrum *and* autocorrelation implementations of the approach is important in order to address recent arguments in favour of alternative methods focusing on phase-locking measures (Rajendran & Schnupp, 2019; Rosso et al., 2021). These methods allow the stability of a phase relationship between two systems (here a periodic beat and the analyzed response) to be addressed directly, as a fundamental way to capture synchronization (also often referred to as "entrainment"; Rosenblum et al., 2001; Pikovsky et al., 2003).

However, this comes with a caveat about how exactly to estimate the phase in the first place. Estimating phase of a system from data is notoriously far from trivial (Kralemann et al., 2007), especially with noisy measurements of complex systems such as the brain (Erra et al., 2017). For example, neuroscientific work has often measured phase values from signals after applying a band-pass filter and computing the Hilbert transform (e.g. Doelling et al., 2019; Rosso et al., 2021), which is equivalent to focusing on a single harmonic of a periodic response in the magnitude spectrum (see also Chang et al., 2022; Ding et al., 2017). As discussed in the current paper, such an approach may overlook instances where the shape of the response may be anything other than a sinusoid. These considerations therefore highlight the advantage of the magnitude spectrum– and autocorrelation-based analyses, which are both sensitive to phase stability without relying on *explicit* phase estimation from data.

6.11 | Conclusion

In conclusion, the frequency-tagging approach in both its magnitude spectrum and new autocorrelation implementations constitutes a valid way to probe beat-related information, in particular beat periodicity, in empirical signals. The novel autocorrelation-based analysis resolves the hurdle posed by the many-to-one relationship between the beat as a perceptual phenomenon, and many equivalent ways in which it can be encoded in time-domain signals captured from the responding system. Likewise, the techniques described here enable rigorous comparison between physical sensory inputs and the responses captured from the system at different processing stages and through different empirical modalities.

Together, these methodological advances open new promising avenues towards a nuanced exploration of processes that enable the mapping between complex sensory inputs such as music and internal representations of time. Specifically, this approach has the potential to facilitate further research into the development, plasticity, and neural basis of this mapping in healthy and impaired populations of human and non-human animals.

7 | METHODS

Simulations and data analyses were carried out using Matlab 9.4.0 (The MathWorks, Natick, MA). Statistical tests were performed using R (4.3.2).

7.1 | Simulations

Responses based on a repeating rhythmic pattern. The signal in Figure 2 was generated based on 11 seamless repetitions of a .2-s-long segment. The segment was created using a method inspired by van Diepen and Mazaheri (2018). The waveform comprised a sum of three sinusoidal components, each generated using the equation below:

$$ERP_{kt} = A_k \frac{t - t_{0,k}}{\tau_k} e^{1-(t-t_{0,k})/\tau_k} \sin(2\pi f_k (t - t_{0,k})) \quad (1)$$

where A_k is the amplitude, f_k is the frequency, t_{0k} is the start time and τ_k is the exponential decay time of the amplitude. The parameters of the three components ($A_1 = .1$, $A_2 = .1$, $A_3 = .2$; $\tau_1 = .2$, $\tau_2 = .05$, $\tau_3 = .3$ s; $f_1 = 1$, $f_2 = 9$, $f_3 = 20$ Hz; all $t_0 s = 0$ s) were chosen arbitrarily, with the intention of creating a smooth yet complex trajectory in the time domain. To decrease the periodic recurrence of the resulting signal, two manipulations were introduced. First, the onset time of each segment was jittered using uniformly distributed values between −15 and +15% of the underlying repetition interval (.2 s). This decreased the phase consistency of the signal at the

rate of the segment repetition. In addition, the shape of each individual segment was slightly changed by adding a snippet of low-pass filtered white noise (Butterworth 4th order filter with 12 Hz cutoff), thus further decreasing the self-similarity of the signal at the segment repetition rate.

The simulations of time-varying responses used for the analyses reported in Figures 3–8 were carried out using the following steps. First, a rhythmic pattern was constructed by arranging several events on a grid of time points separated by a fixed interval (here .2 s). Unless stated otherwise, the pattern used across simulations in the current paper could be represented as [x.xxxx.xxx..], where "x" stands for a grid position with an event and "." marks an empty grid position. These specific rhythmic pattern and grid interval were intentionally identical to the "weakly periodic rhythm" used as a stimulus in the empirical datasets (Lenc et al., 2018, 2022), for comparison purposes. The groups of events in the pattern were arranged in a way that did not clearly align with any plausible periodic beat. Therefore, time-varying signals based on this rhythmic pattern were expected to exhibit low periodic recurrence at different beat rates allowed by the underlying grid structure. The 2.4-s-long pattern (12 × .2 s) was then seamlessly repeated 20 times to form a long sequence.

Furthermore, a time-domain signal was generated by assigning a unitary response to each grid point that contained an event. The unitary response could be either a Dirac impulse (e.g. Figure S1B), a square wave (e.g. Figures 3 and 5) or a simulated event-related potential (ERP; Luck, 2014) (see Figures S1C and 4). The ERP was generated using a method inspired by van Diepen and Mazaheri (2018). The waveform comprised a sum of several sinusoidal components generated using the equation below:

$$ERP_{kt} = A_k \frac{t - t_{0,k}}{\tau_k} e^{1-(t-t_{0,k})/\tau_k} \sin(2\pi f_k (t - t_{0,k})) \quad (2)$$

where A_k is the amplitude, f_k is the frequency, t_{0k} is the start time and τ_k is the exponential decay time of the amplitude. Two frequency components were summed to generate the complex ERP-like unitary response with total duration restricted to .5 s. The first component corresponded to a fast transient response, akin to the P50 evoked potential (Picton, 2010) ($A_1 = .75$, $f_1 = 7$ Hz, $t_{0,1} = 0$ s, $\tau_1 = .05$ s). The second component had a longer integration time and contributed to the slow dynamics of the response ($A_2 = .4$, $f_2 = 1$ Hz, $t_{0,2} = 0$ s, $\tau_2 = .2$ s).

Periodic recurrence of the signal at the rate of 1/.8 s = 1.25 Hz (corresponding to a beat period spanning four grid points) was selectively enhanced by manipulating the amplitude of the unitary response assigned to each grid position. In particular, the amplitude of unitary responses that occurred at grid points overlapping with integer multiples of the beat period (i.e. .8 s, 1.6 s, 3.2 s, etc.) was kept constant, while the amplitude of all other unitary responses was reduced. The beat rate of 1.25 Hz was chosen to match the most consistently perceived beat rate in the empirical studies re-analyzed in the current paper (see Section 5.3). Likewise, the selection of beat-related and -unrelated frequencies and lags was equivalent to the selection used to analyze the EEG data, except of simulations for Figure 6 that were carried out using a single target lag set to .8 s and flanking lags .6, 1.0 and 1.2 s chosen for simplicity to demonstrate the effect of standardization.

To assess the sensitivity of the magnitude spectrum-based and autocorrelation-based analyses to phase stability, a perfectly periodic signal was generated at the beat rate of 1.25 Hz using a repeated square-wave unitary response. All other parameters were identical to the sequences described above. Periodic recurrence was then systematically decreased by randomly jittering the onset times of each individual unitary response (normal distribution with standard deviation log-spaced between 10 and 200 ms across conditions), as shown in Figure S5.

To illustrate the magnitude spectrum and autocorrelation function computed from a non-repeating rhythmic signal constructed on an isochronous grid of time points (Figure S7), square-wave unitary responses were assigned to positions on a 240-point grid (.2 s grid interval) to form a long sequence. Whether or not a grid position would contain a unitary response was determined at random, with the constraint that a response occurred at more than half of grid points within the sequence, which corresponded to integer multiples of a .8-s period, in order to slightly enhance the periodic recurrence at the rate of 1/.8 s = 1.25 Hz.

7.1.1 | Signal transformations non-specific to periodicity

In order to probe the specificity of a method to capturing periodic recurrence, several transformations that preserve periodicity were applied to the simulated signals. Firstly, multiplication by a constant was used to change the *scale* of the signal. Secondly, adding a constant offset to the signal was used to introduce a *shift*. Finally, the *shape* of the signal was changed in a way that does not affect periodic recurrence at a given rate. This was done in the context of signals based on a grid of time points and a repeating rhythmic pattern as described above. Specifically, the shape of the unitary response was changed, while

keeping all other parameters constant. Several versions of the same signal were generated using square-wave unitary responses with duty cycles (i.e. the proportion of the grid interval during which the square wave has a "high" vs. "low" value) ranging from 25% to 90%.

7.1.2 | Simulating noisy signals

The noise used across all simulations was sampled from an open dataset of resting-state EEG recordings (Wang et al., 2022). The dataset comprises brain activity of 60 participants, recorded at 500 Hz sampling rate using 64 active Ag-AgCl electrodes placed according to the 10/20 international system (Brain Products GmbH, Germany). The study was approved by the Review Board of the Institute of Southwest University, and written informed consent was obtained from all the participants.

Resting-state EEG data from the "Eyes Open" task were used. These data consisted of 5-min recordings where participants were instructed to fixate a point in front of them and avoid any unnecessary movement. The EEG data were re-referenced to common average. The noise for each individual simulated signal was prepared by (1) selecting a random participant, (2) selecting a random channel, (3) resampling to the required sampling rate and (4) randomly picking segment of the required duration within the 5-min recording. Finally, the sampled noise segment was summed with a simulated response to yield a noisy signal. The response and noise were scaled to yield a desired SNR defined by the equation below:

$$SNR = 20 * \log 10\left[\frac{rms(x_{signal})}{rms(x_{noise})}\right] \quad (3)$$

where $rms(x_{signal})$ is the root-mean-square amplitude of the simulated response and $rms(x_{noise})$ is the root-mean-square amplitude of the noise segment.

To investigate how the noise level affected the mean z-scored autocorrelation at beat-related lags, two example response signals were generated. Both responses were based on a repeating rhythmic pattern generated using an isochronous grid of time points, as described above. The pattern for each response was chosen from a set of all possible seven-event patterns constructed on a 12-point grid with a restriction that at most four successive grid points could contain an event.

For each pattern, a time-varying response was generated using a square wave as a unitary response. The mean z-scored autocorrelation at beat-related lags was obtained from each simulated response. The response with beat-related z-score closest to .5 (based on the pattern [xx.xx.x.x.x.]) was selected as an example of a large positive z-score. The second response (based on the pattern [xx.xx.x.xx..]) was selected since its beat-related z-score was closest to −.5, thus serving as an example of a largely negative z-score.

The two selected example responses were used to generate noisy signals with 10 different SNR levels linearly between −40 and 18 dB (50 simulations generated for each SNR level). The z-score at beat-related lags was estimated from the autocorrelation obtained either directly from each simulated noisy signal or after applying the full noise-correction method described in Sections 5.2 and 7.3. In addition, the same analysis was applied to 500 signals generated only from noise (i.e. without any response).

In order to investigate how the noise distorts the values obtained from the autocorrelation function, a range of noisy signals were generated. The signals were based on responses generated from five randomly chosen rhythmic patterns ([xx.xxx.x.x..], [xxxx.x.x.x..], [xxx.x.xx.x..], [xxxx.xx.x...], [xxx.xx.x.x..]). For each pattern, 50 noisy signals were prepared separately for 10 SNR levels (linearly between −40 and 18 dB). All other parameters were identical to the simulations and analyses described above. The empirical zSNR was estimated from the spectrum of each simulated signal by pooling all beat-related and -unrelated frequencies and taking adjacent frequency bin 2 to 5 from each side to estimate the local noise baseline (see Section 7.3 for details).

The autocorrelation function was estimated either (i) directly from the noisy signal, (ii) after estimating and subtracting the 1/f noise component using IRASA or (iii) by additionally zeroing-out magnitudes at frequencies that did not capture the simulated response (see Section 7.3 for details). The autocorrelation values extracted across all beat-related and -unrelated lags were arranged into a single vector. For each signal, this vector was correlated (Pearson's r) with the equivalent vector obtained from the autocorrelation of the corresponding response without noise (i.e. the "ground-truth"). This correlation was used to quantify how well the response autocorrelation can be reconstructed from the noisy signal. Finally, the distribution density of the obtained correlation coefficients was estimated separately for 11 equally spaced bins covering the empirical zSNR range of all simulated signals.

7.2 | Analysis

7.2.1 | DFT

The DFT of a time-domain signal was calculated using the *fft* function in Matlab, which returns an array of

complex-valued coefficients, one per frequency bin. Absolute value of each complex coefficient was taken to obtain the magnitude spectrum of the signal.

7.2.2 | Autocorrelation

The circular autocorrelation of the time-domain signal was efficiently computed in the frequency domain, using the convolution theorem (Smith, 2007) as shown in the equation below

$$ACF_t = Re(ifft(fft(x_t)\,fft(x_t)^*)) \quad (4)$$

where ACF_t corresponds to the autocorrelation (as a function of time), x_t is the time-domain signal, *fft* is the DFT (implemented via Fast Fourier Transform), *ifft* is the inverse DFT, *Re* indicates taking the real component of a complex number and $*$ represents complex conjugation. Specifically, the complex-valued DFT of the signal was point-wise multiplied by its complex conjugate and transformed back to the time-domain using the inverse DFT followed by taking the real component.

7.2.3 | Standardized metrics of relative prominence

The main goal of standardization in the context of the current paper is to obtain a metric which quantifies the prominence of magnitudes at a set of target frequencies (here beat-related frequencies), relative to a set of standardization frequencies (here beat-unrelated frequencies). This relative prominence can be measured by calculating, for example, a ratio or contrast using the following equations:

$$ratio_{target\ freqs} = \frac{mean\left(X_{target\ freqs}\right)}{mean\left(X_{standardization\ freqs}\right)} \quad (5)$$

$$contrast_{target\ freqs} = \frac{mean\left(X_{target\ freqs}\right) - mean\left(X_{standardization\ freqs}\right)}{mean\left(X_{target\ freqs}\right) + mean\left(X_{standardization\ freqs}\right)} \quad (6)$$

Alternatively, each value can be *z*-scored using the equation below:

$$z\,score_i = \frac{X_i - mean\left(X_{all\ freqs}\right)}{sd\left(X_{all\ freqs}\right)} \quad (7)$$

where $z\,score_i$ is the *z*-scored magnitude at frequency i, X_i is the raw magnitude at frequency i (X_i being the DFT of the time-domain signal x_t), $X_{all\ freqs}$ indicates a set of magnitudes at all frequencies of interest (i.e. all target and standardization frequencies), *mean* indicates an average and *sd* corresponds to the standard deviation. The obtained *z*-scored magnitudes at a set of target frequencies can be averaged to calculate a composite measure of their relative prominence.

The exact same metrics can be used to quantify the relative prominence of autocorrelation values at a set of chosen target lags.

7.3 | Accounting for noise

7.3.1 | Accounting for noise in DFT

The frequency-tagging approach relies on the ability to reliably measure the magnitudes at several frequencies of interest in the spectrum of the response. However, magnitudes taken from the spectrum of a raw recorded signal can be significantly biased by noise, as discussed in Section 5.2. In order to minimize the contribution of noise to the estimated magnitudes, a common approach capitalizes on the fact that the noise is expected to have a smooth and broadband spectrum, while the spectrum of a periodic response only contains narrow peaks at a-priori known frequency bins (see Section 4). Hence, the noise-corrected magnitude at a frequency bin of interest can be approximated by measuring how much the peak at that frequency stands out relative to the local noise baseline. This can be simply quantified by taking the difference between the magnitude at the bin of interest and the average magnitude at adjacent frequency bins (see Figure S4).

A similar rationale can be used to quantify how much the response of interest stands out from the noise, a quantity often referred to as "SNR" (Meigen & Bach, 2000). As shown in Figure S4, the magnitude and local noise baseline can be pooled (i.e. summed) across all response frequencies (i.e. frequencies corresponding to the rate of the periodic response and harmonics). Subsequently, a *z*-scored signal-to-noise ratio (zSNR) can be calculated using the following equation:

$$zSNR_{response\ bin} = \frac{X_{response\ bin} - mean\left(X_{adjacent\ bins}\right)}{sd\left(X_{adjacent\ bins}\right)} \quad (8)$$

where $X_{response\ bin}$ is the magnitude summed across response frequencies, $mean\left(X_{adjacent\ bins}\right)$ is the average magnitude of the neighboring bins summed across response frequencies (i.e. local noise baseline) and $sd\left(X_{adjacent\ bins}\right)$ is the standard deviation across the

neighboring bins. This approach has been widely used in the frequency-tagging literature, as it offers a straightforward way to statistically test whether the response stands out significantly from the noise in the recorded signal (Hagen et al., 2021; Jonas et al., 2016; Liu-Shuang et al., 2014; Lochy et al., 2018; Volfart et al., 2020).

7.3.2 | Accounting for noise in autocorrelation

In order to minimize the impact of noise on the autocorrelation-based estimate of beat periodicity, we used a two-step procedure.

First, the 1/f-like noise magnitude was estimated using IRASA by irregular resampling of the signal in the time domain (Wen & Liu, 2016). The advantage of using this method is that it does not rely on fitting an explicit function to the magnitude spectrum (cf., e.g. Donoghue et al., 2020). Rather, it capitalizes on the fact that the magnitude spectrum of an aperiodic 1/f signal (here considered noise) is invariant when the signal is resampled, whereas the spectrum of a periodic signal (here considered response of interest) will be strongly affected by resampling. The method was implemented using a Matlab library published in Wen and Liu (2016). Nineteen scaling factors were used, spaced between 1.1 and 2 in equal steps of .05. As noted by Gerster et al. (2022), the resampling procedure causes a reduction of bandwidth determined by the highest scaling factor (here 2). However, the original sampling rate of the simulated and empirical signals used in the current study was always sufficiently high to cover the frequency range of interest even after a bandwidth reduction by a factor of 2.

Using these parameters, IRASA estimated the magnitude spectrum of the aperiodic noise component from 0 Hz up to 1/4 of the original sampling rate. The part of the estimated noise spectrum that was missing due to bandwidth reduction (i.e. from 1/4 to 1/2 of the sampling rate) was filled with zeros. Finally, the full magnitude spectrum of the noise component was reconstructed by mirroring around 1/2 of the sampling rate (i.e. the Nyquist frequency), using the conjugate symmetry of the DFT for real signals (Smith, 2007).

The magnitude of the estimated noise spectrum was then subtracted from the complex-valued spectrum of the raw signal, separately for each frequency bin, as follows. If the magnitude of the noise was larger than the magnitude of the signal, the value at the corresponding frequency was set to a zero vector. Otherwise, the length of the complex vector in signal's spectrum was reduced by the amount equal to the magnitude of the noise at the corresponding frequency.

After the noise subtraction, the second step comprised further suppressing the effect of noise by only keeping frequencies that captured the response of interest. The selection of these frequencies depends on the particular design and goals of the analysis and is further discussed in Section 5.2. The complex coefficient at any frequency bin that did not overlap with the chosen set of response frequencies was set to a zero vector. This "zeroing" operation is justified by the fact that the response was elicited by a stimulus made up of a seamlessly looped rhythmic pattern, hence was expected to itself periodically repeat at the same rate as the rhythmic pattern in the input (in fact, the assumption that the same response will be consistently elicited when repeating the same stimulus or experimental condition constitutes a fundamental principle used in neurosciences to isolate a response from noise). In addition, the design of the stimulus sequence ensured that an exact integer number of pattern-repetition periods was captured. Consequently, the DFT of any response obtained with the paradigm used here can be expected to only contain energy at the exact frequency bins corresponding to integer multiples of the rhythmic pattern repetition rate.

Notably, this set of frequencies comprises all beat-related and -unrelated frequencies (as discussed in Section 4.3); thus, zeroing out frequency bins outside of this set is not expected to systematically bias the periodic recurrence of the signal at the rate of the beat targeted in the current analysis. Nevertheless, it is important to note that the "zeroing" operation will yield a periodic signal when converted back to the time domain (similarly to what a very sharp comb filter would do). In other words, only keeping frequencies corresponding to harmonics of the pattern repetition rate will result in a signal that periodically repeats at the rate of the pattern repetition. Critically, the shape of the repeated signal segment will be determined by the characteristics of the original signal, and thus, it is valid to analyze this shape to learn about the prominence of periodicities nested within the rhythmic pattern, such as the beat periodicity measured in our analysis. In saying that, there are several remarks that deserve attention.

The fact that the noise "zeroing" step makes the signal periodic in the time domain means that the ACF of the signal will be likewise periodic, strictly repeating itself at the rate determined by the duration of the rhythmic pattern. Hence, the ACF values taken from lags higher than the pattern duration (in fact, half of the pattern duration due to the symmetries in ACF) will be redundant. Why then consider lags all the way to half signal duration, as done in the current analysis (see Section 5.1)? One reason is that this is a straightforward way to weight the lags according to how much they

capture the different periodicities taken as beat-related, as well as beat-unrelated. Specifically, here, the beat period of .8 s was evaluated relative to the prominence of periods equal to .6, 1.0 and 1.4 s, which correspond to plausible alternative periodic pulses that do not necessarily complete an integer number of cycles within the duration of the repeated rhythmic pattern (here 2.4 s). Hence, these beat-unrelated periodicities will project to lags beyond 2.4 s in the ACF of the original signal and will therefore contribute differently to the shape of the periodically repeating ACF segment after the noise "zeroing" step has been applied. This can be simply accounted for by considering multiples of the beat-unrelated periods all the way up to half signal duration.

Relatedly, the periodicity in the ACF after applying the noise "zeroing" step may lead to increased variance of the observed *z*-scores at beat-related lags when the noise level is high. This is because having only a few unique autocorrelation values in the periodically repeating ACF segment increases the likelihood that very high or very low mean *z*-score at beat-related lags will be observed by chance, if the signal is dominated by random noise. This can be observed in Figure 7b,c, whereby the correlation between ground-truth ACF values and values reconstructed from a noisy version of the signal become uniformly distributed between −1 and 1 across many simulations, indicating the instability of the estimate once noise level is too high.

This high variance can be reduced by leaving a small frequency band around each response frequency that smoothly tappers off, rather than abruptly setting all frequency bins other than the response bins to zero. To illustrate this effect (see Figure S9), a small symmetrical band was retained around each response frequency, with a magnitude linearly decreasing from 1 to 0 starting from the half of the band on each side of the response frequency bin. Indeed, as shown in Figure S9, allowing a wider band around each response frequency bin reduces the variance of similarity between ground-truth ACF values and the estimated ACF values at high noise levels, as the noise is allowed to yield unique random ACF values all the way up to the lag corresponding to the half of signal duration. This would correspondingly decrease the variance of *z*-score values estimated from such high-noise signals. However, Figure S9 also shows that leaving a small band around each response frequency decreases the ability to reconstruct ground truth ACF at medium noise levels. This can be thought of as a form of bias-variance trade-off, and choosing the width of the band taken around each response frequency should be informed by the knowledge of the empirical data, as well as the particular goals of the analysis.

7.4 | Empirical data

Datasets from two previously published studies (Lenc et al., 2018, 2022) were used. A detailed description of the experimental design and procedures, as well as the data acquisition parameters and preprocessing are provided in the original publications.

7.4.1 | Stimuli

The stimuli consisted of rhythmic auditory sequences. Each sequence was created by seamlessly looping a 2.4-s-long rhythmic pattern. The pattern was generated by arranging eight identical pure tones on an isochronous grid of 12 time points separated by .2 s. The rhythm was either made up of *low-pitched tones* (130 Hz) or *high-pitched tones* (1236.8 Hz). In one condition, the constituent pattern was a "strongly periodic rhythm", since the groups of tones were arranged in a way that closely matched a beat with a rate of 1.25 Hz. In the other condition, the sequence was made of a repeating "weakly periodic rhythm", whereby the tones were arranged in a way that did not systematically match any plausible periodic beat.

7.4.2 | Cochlear model

A biomimetic model was used to obtain a lower level sensory representation of the stimulus (Slaney, 1998). This model (hereafter cochlear model) simulates cochlear filtering and subsequent nonlinearities due to neural transduction in the inner hair-cell/auditory-nerve synapse. The time-domain output of the model can be considered as an approximation of the mean firing rate response in the auditory nerve, that is, at an early subcortical stage of the auditory pathway.

7.4.3 | Adult dataset

The EEG activity was recorded from 14 healthy adult participants using a Biosemi Active-Two system (Biosemi, Amsterdan, The Netherlands) with 64 Ag-AgCl placed on the scalp according to the international 10/20 system. All participants gave informed consent, and the study was approved by the Research Ethics Committee of Western Sydney University. Participants were asked to listen to auditory sequences in 50.4-second-long trials (eight trials per condition), avoid unnecessary movement and carry out a temporal deviant identification task to ensure their attention to the stimuli.

The EEG data were re-referenced to the common average electrode, further low-pass filtered at 30 Hz (2nd order Butterworth filter), and averaged across nine frontocentral channels and trials. The nine channels were chosen since high SNR of responses to rhythmic stimuli has been consistently observed at these channels throughout prior studies (Kaneshiro et al., 2020; Lenc et al., 2020; Nozaradan et al., 2012). Transforming the data into the frequency domain using DFT yielded a spectrum with frequency resolution of 1/50.4 s = .02 Hz. The limit up to which unique autocorrelation values could be obtained from the EEG responses was equal to half the trial duration, i.e., 50.4/2 = 25.2 s.

7.4.4 | Infant dataset

The EEG activity was recorded from 20 infants (aged from 5 to 6 months) using a 128-channel HydroCel GSN net and an Electrical Geodesic NetAmps 200 amplifier. The research was approved by the Research Ethics Committee of Western Sydney University, as part of the project H9660, and informed consent was obtained from legal representatives of the infants. The infants were passively listening to the rhythmic stimuli in 60-s-long trials (five trials per condition). The pre-processed EEG data were re-referenced to the average of mastoid electrodes and averaged across 28 frontocentral channels. Data from each trial were further cut into two 26.4-s epochs, and these were averaged within and across trials. Since EEG of young infants is typically highly contaminated with low-frequency artifacts, the data were filtered with a relatively high cut-off frequency during preprocessing (.5 Hz). The cochlear model output was filtered using the same parameters to ensure valid comparison with the EEG responses. Because the filtering strongly affected the lowest frequency of interest at .416 Hz (i.e. the first harmonic of the stimulus pattern repetition rate), this frequency was excluded from further analyses. Given the pre-processed response had a duration of 26.4 s, the DFT yielded a spectrum with frequency bins separated by 1/26.4 s = .038 Hz. Likewise, the autocorrelation function yielded unique values up the lag of 26.4/2 s.

7.4.5 | Continuous tapping data

The dataset of Lenc et al. (2018) also contained tapping data collected from each participant after the EEG session. Participants tapped their finger on a custom-built box with a piezoelectric sensor, which captured the mechanical vibrations elicited by the impact of the tapping finger. The resulting signal contained information about the intensity or force with which each tap was executed as a function of time. The continuous tapping responses were analyzed separately for each trial (two trials per participant and condition), otherwise using the same procedure and parameters that were applied to the EEG data.

AUTHOR CONTRIBUTIONS

T. L., C. L., and S. N. designed the study, did the analyses and created the figures. All authors contributed to writing and editing the paper.

ACKNOWLEDGEMENTS

T. L. receives funding from the HORIZON EUROPE Marie Skłodowska-Curie programme (101148958). S. N. is supported by a Starting Grant from the H2020 European Research Council (801872). P. E. K. has support from the Danish National Research Foundation (DNRF117). D. M. is a research fellow of the Fonds de la Recherche Scientifique - FNRS (FC17797).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ejn.16637.

DATA AVAILABILITY STATEMENT

Project repository: The data and code that support the findings of this study are openly available in at https://osf.io/5s3j7, reference number 5s3j7. Datasets: All empirical data analyzed in this manuscript are secondary uses of data that have been previously published and/or were accessed from openly available data repositories. The empirical datasets, as well as the code necessary to re-create the simulated datasets, are included in the project repository. The resting-state EEG recordings used to obtain noise samples are part of an open dataset available in OpenNEURO at https://openneuro.org/datasets/ds004148/versions/1.0.1, accession number ds004148. Software: Code for this project was written in MATLAB and is deposited in the project repository. In addition, a MATLAB library to perform the autocorrelation analysis developed in the current project is openly available on github (https://github.com/TomasLenc/acf_tools) and licensed for reuse.

ORCID

Tomas Lenc https://orcid.org/0000-0001-5796-1388
Cédric Lenoir https://orcid.org/0000-0002-1420-7550
Peter E. Keller https://orcid.org/0000-0001-7579-6515

Rainer Polak https://orcid.org/0000-0001-5684-1499
Dounia Mulders https://orcid.org/0000-0003-4855-5331
Sylvie Nozaradan https://orcid.org/0000-0002-5662-3173

REFERENCES

Ahrens, M. B., Linden, J. F., & Sahani, M. (2008). Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. *Journal of Neuroscience*, *28*, 1929–1942. https://doi.org/10.1523/JNEUROSCI.3377-07.2008

Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition*, *48*, 66–79. https://doi.org/10.1006/brcg.2001.1304

Asokan, M. M., Williamson, R. S., Hancock, K. E., & Polley, D. B. (2021). Inverted central auditory hierarchies for encoding local intervals and global temporal patterns. *Current Biology*, *31*, 1762–1770.e4. https://doi.org/10.1016/j.cub.2021.01.076

Bach, M., & Meigen, T. (1999). Do's and don'ts in Fourier analysis of steady-state potentials. *Documenta Ophthalmologica*, *99*, 69–82. https://doi.org/10.1023/A:1002648202420

Bamford, J. S., Burger, B., & Toiviainen, P. (2023). Turning heads on the dance floor: Synchrony and social interaction using a silent disco paradigm. *Music & Science*, *6*, 20592043231155416. https://doi.org/10.1177/20592043231155416

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. *Psychophysiology*, *55*, e13049. https://doi.org/10.1111/psyp.13049

Bouvet, C. J., Bardy, B. G., Keller, P. E., Bella, S. D., Nozaradan, S., & Varlet, M. (2020). Accent-induced modulation of neural and movement patterns during spontaneous synchronization to auditory rhythms. *Journal of Cognitive Neuroscience*, *32*, 2260–2271. https://doi.org/10.1162/jocn_a_01605

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "Ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14*, 362–366. https://doi.org/10.1111/1467-9280.24441

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid transformation from auditory to linguistic representations of continuous speech. *Current Biology*, *28*, 3976–3983.e5. https://doi.org/10.1016/j.cub.2018.10.042

Brown, J. C. (1993). Determination of the meter of musical scores by autocorrelation. *Journal of the Acoustical Society of America*, *94*, 1953–1957. https://doi.org/10.1121/1.407518

Butler, M. J. (2006). *Unlocking the groove: Rhythm, meter, and musical design in electronic dance music*. Indiana University Press.

Câmara, G. S., & Danielsen, A. (2018). Groove. In A. Rehding & S. Rings (Eds.), *Oxford handbook of critical concepts in music theory* (pp. 271–294). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190454746.013.17

Cameron, D. J., Zioga, I., Lindsen, J. P., Pearce, M. T., Wiggins, G. A., Potter, K., & Bhattacharya, J. (2019). Neural entrainment is associated with subjective groove and complexity for performed but not mechanical musical rhythms. *Experimental Brain Research*, *237*, 1981–1991. https://doi.org/10.1007/s00221-019-05557-4

Cannon, J. (2021). Expectancy-based rhythmic entrainment as continuous Bayesian inference. *PLoS Computational Biology*, *17*, e1009025. https://doi.org/10.1371/journal.pcbi.1009025

Celma-Miralles, A., de Menezes, R. F., & Toro, J. M. (2016). Look at the beat, feel the meter: Top-down effects of meter induction on auditory and visual modalities. *Frontiers in Human Neuroscience*, *10*, 108. https://doi.org/10.3389/fnhum.2016.00108

Celma-Miralles, A., & Toro, J. M. (2019). Ternary meter from spatial sounds: Differences in neural entrainment between musicians and non-musicians. *Brain and Cognition*, *136*, 103594. https://doi.org/10.1016/j.bandc.2019.103594

Chang, E. F. (2015). Towards large-scale, human-based, mesoscopic neurotechnologies. *Neuron*, *86*, 68–78. https://doi.org/10.1016/j.neuron.2015.03.037

Chang, A., Teng, X., Assaneo, F., & Poeppel, D. (2022). The human auditory system uses amplitude modulation to distinguish music from speech. *PLOS Biology*, *22*, e3002631.

Chemin, B., Huang, G., Mulders, D., & Mouraux, A. (2018). EEG time-warping to study non-strictly-periodic EEG signals related to the production of rhythmic movements. *Journal of Neuroscience Methods*, *308*, 106–115. https://doi.org/10.1016/j.jneumeth.2018.07.016

Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body movement selectively shapes the neural representation of musical rhythms. *Psychological Science*, *25*, 2147–2159. https://doi.org/10.1177/0956797614551161

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring neural entrainment to beat and meter in infants: Effects of music background. *Frontiers in Neuroscience*, *10*, 229. https://doi.org/10.3389/fnins.2016.00229

Clarke, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 19–33). Royal Swedish Academy of Music.

Cohn, R. (2014). Meter without Tactus. In J. Lochhead & R. Will (Eds.), *Society for music theory annual meeting*. American Musicological Society.

Cohn, R. (2020). Meter. In A. Rehding & S. Rings (Eds.), *Oxford handbook of critical concepts in music theory* (pp. 207–233). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190454746.013.9

Danielsen, A., Johansson, M., & Stover, C. (2023). Bins, spans, and tolerance: Three theories of microtiming behavior. *Music Theory Spectrum*, *45*, 181–198. https://doi.org/10.1093/mts/mtad005

Daube, C., Ince, R. A. A., & Gross, J. (2019). Simple acoustic features can explain phoneme-based predictions of cortical responses to speech. *Current Biology*, *29*, 1924–1937. https://doi.org/10.1016/j.cub.2019.04.067

David, S. V., Mesgarani, N., Fritz, J. B., & Shamma, S. A. (2009). Rapid synaptic depression explains nonlinear modulation of spectro-temporal tuning in primary auditory cortex by natural stimuli. *Journal of Neuroscience*, *29*, 3374–3386. https://doi.org/10.1523/JNEUROSCI.5249-08.2009

van Diepen, R. M., & Mazaheri, A. (2017). Cross-sensory modulation of alpha oscillatory activity: Suppression, idling, and default resource allocation. *European Journal of Neuroscience*, *45*, 1431–1438. https://doi.org/10.1111/ejn.13570

van Diepen, R. M., & Mazaheri, A. (2018). The caveats of observing inter-trial phase-coherence in cognitive neuroscience. *Scientific Reports*, *8*, 2990. https://doi.org/10.1038/s41598-018-20423-z

van Diepen, R. M., Miller, L. M., Mazaheri, A., & Geng, J. J. (2016). The role of alpha activity in spatial and feature-based attention. *eNeuro*, *3*, e0204–16.2016. https://doi.org/10.1523/ENEURO.0204-16.2016

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*, *81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences*, *116*, 10113–10121. https://doi.org/10.1073/pnas.1816414116

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience*, *23*, 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Drennan, D. P., & Lalor, E. C. (2019). Cortical tracking of complex sound envelopes: Modeling the changes in response with intensity. *eNeuro*, *6*, 1–11. https://doi.org/10.1523/ENEURO.0082-19.2019

Erra, R. G., Velazquez, J. L. P., & Rosenblum, M. (2017). Neural synchronization from the perspective of non-linear dynamics. *Frontiers in Computational Neuroscience*, *11*, 1–4.

Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, *35*, 15187–15198. https://doi.org/10.1523/JNEUROSCI.2397-15.2015

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology*, *17*, e3000054. https://doi.org/10.1371/journal.pbio.3000054

Gerster, M., Waterstraat, G., Litvak, V., Lehnertz, K., Schnitzler, A., Florin, E., Curio, G., & Nikulin, V. (2022). Separating neural oscillations from aperiodic 1/f activity: Challenges and recommendations. *Neuroinformatics*, *20*, 991–1012. https://doi.org/10.1007/s12021-022-09581-8

Goldstone, R. L., Kersten, A., & Carvalho, P. F. (2018). Categorization and concepts. In *Stevens' handbook of experimental psychology and cognitive neuroscience* (pp. 1–43). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119170174.epcn308

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*, 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Grootswagers, T., McKay, H., & Varlet, M. (2022). Unique contributions of perceptual and conceptual humanness to object representations in the human brain. *NeuroImage*, *257*, 119350. https://doi.org/10.1016/j.neuroimage.2022.119350

Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage*, *188*, 668–679. https://doi.org/10.1016/j.neuroimage.2018.12.046

Groppe, D. M., Bickel, S., Keller, C. J., Jain, S. K., Hwang, S. T., Harden, C., & Mehta, A. D. (2013). Dominant frequencies of resting human brain activity as measured by the electrocorticogram. *NeuroImage*, *79*, 223–233. https://doi.org/10.1016/j.neuroimage.2013.04.044

Hagen, S., Jacques, C., Maillard, L., Colnat-Coulbois, S., Rossion, B., & Jonas, J. (2020). Spatially dissociated intracerebral maps for face- and house-selective activity in the human ventral occipito-temporal cortex. *Cerebral Cortex*, *30*, 4026–4043. https://doi.org/10.1093/cercor/bhaa022

Hagen, S., Lochy, A., Jacques, C., Maillard, L., Colnat-Coulbois, S., Jonas, J., & Rossion, B. (2021). Dissociated face- and word-selective intracerebral responses in the human ventral occipito-temporal cortex. *Brain Structure and Function*, *226*, 3031–3049. https://doi.org/10.1007/s00429-021-02350-4

Hamilton, L. S., Oganian, Y., Hall, J., Chang, E. F., Francisco, S., & Sciences, H. (2021). Parallel and distributed encoding of speech across human auditory cortex. *Cell*, *184*, 1–14. https://doi.org/10.1016/j.cell.2021.07.019

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 956–974.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*, 466–472. https://doi.org/10.1016/j.tics.2007.08.008

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 12639–12643. https://doi.org/10.1073/pnas.0504254102

Harrison, P. M. C., Bianco, R., Chait, M., & Pearce, M. T. (2020). PPM-decay: A computational model of auditory prediction with memory decay. *PLoS Computational Biology*, *16*, e1008995. https://doi.org/10.1371/journal.pcbi.1008995

He, B. J. (2014). Scale-free brain activity: Past, present, and future. *Trends in Cognitive Sciences*, *18*, 480–487. https://doi.org/10.1016/j.tics.2014.04.003

He, B. J., Zempel, J. M., Snyder, A. Z., & Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron*, *66*, 353–369. https://doi.org/10.1016/j.neuron.2010.04.020

Herff, S. A., Herff, C., Milne, A. J., Johnson, G. D., Shih, J. J., & Krusienski, D. J. (2020). Prefrontal high gamma in ECoG tags periodicity of musical rhythms in perception and imagination. *eNeuro*, *7*, 1–11.

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, and Psychophysics*, *72*, 1218–1227. https://doi.org/10.3758/APP.72.5.1218

Honing, H., & Bouwer, F. L. (2018). Rhythm. In P. J. Rentfrow & D. J. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 33–70). MIT Press.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, *1169*, 58–73. https://doi.org/10.1111/j.1749-6632.2009.04579.x

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, *27*, 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Jacoby N, Polak R., Grahn J.A., Cameron D.J., Lee K.M., Godoy R., Undurraga E.A., Huanca T., Thalwitzer T., Doumbia N., Goldberg D., Margulis E.H., Wong P.C.M., Jure L., Rocamora M., Fujii S., Savage P.E., Ajimi J., Konno R., ..., McDermott J.H. (2024) Commonality and variation in mental representations of music revealed by a cross-cultural comparison of rhythm priors in 15 countries. *Nature Human Behaviour* Available at: https://www.nature.com/articles/s41562-023-01800-9 [Accessed March 5, 2024], *8*, 846, 877, https://doi.org/10.1038/s41562-023-01800-9

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., & Rossion, B. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proceedings of the National Academy of Sciences*, *113*, E4088–E4097. https://doi.org/10.1073/pnas.1522033113

Jones, M. R., & Mcauley, J. D. (2005). Time judgments in global temporal contexts. *Perception & Psychophysics*, *67*, 398–417. https://doi.org/10.3758/BF03193320

Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., & Berger, J. (2020). Natural music evokes correlated EEG responses reflecting temporal structure and beat. *NeuroImage*, *214*, 116559. https://doi.org/10.1016/j.neuroimage.2020.116559

Kappenman, E. S., & Luck, S. J. (2011). ERP components: The ups and downs of brainwave recordings. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components* (pp. 3–30). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Keil, A., Bernat, E. M., Cohen, M. X., Ding, M., Fabiani, M., Gratton, G., Kappenman, E. S., Maris, E., Mathewson, K. E., Ward, R. T., & Weisz, N. (2022). Recommendations and publication guidelines for studies using frequency domain and time-frequency domain analyses of neural time series. *Psychophysiology*, *59*, e14052. https://doi.org/10.1111/psyp.14052

Keller, P., & Schubert, E. (2011). Cognitive and affective judgements of syncopated musical themes. *Advances in Cognitive Psychology*, *7*, 142–156. https://doi.org/10.2478/v10053-008-0094-0

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, *93*, 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Kralemann, B., Cimponeriu, L., Rosenblum, M., Pikovsky, A., & Mrowka, R. (2007). Uncovering interaction of coupled oscillators from data. *Physical Review E*, *76*, 055201. https://doi.org/10.1103/PhysRevE.76.055201

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*, 401–412. https://doi.org/10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4. https://doi.org/10.3389/neuro.06.004.2008

Kriegeskorte, N., & Wei, X.-X. (2021). Neural tuning and representational geometry. *Nature Reviews Neuroscience*, *22*, 703–718. https://doi.org/10.1038/s41583-021-00502-3

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time* (pp. 189–231). Emerald Publishing Limited.

Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research*, *66*, 3–17. https://doi.org/10.1007/s004260100069

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, *9*, 159. https://doi.org/10.3389/fnsys.2015.00159

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37. https://doi.org/10.1207/s15516709cog2601_1

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the new York Academy of Sciences*, *1169*, 46–57. https://doi.org/10.1111/j.1749-6632.2009.04550.x

Leman, M., & Naveda, L. (2010). Basic gestures as spatiotemporal reference frames for repetitive dance/music patterns in samba and Charleston. *Music Perception*, *28*, 71–91. https://doi.org/10.1525/mp.2010.28.1.71

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2018). Neural tracking of the musical beat is enhanced by low-frequency sounds. *Proceedings of the National Academy of Sciences*, *115*, 8221–8226. https://doi.org/10.1073/pnas.1801421115

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2019). Reply to Rajendran and Schnupp: Frequency tagging is sensitive to the temporal structure of signals. *Proceedings of the National Academy of Sciences*, *116*, 2781–2782. https://doi.org/10.1073/pnas.1820941116

Lenc, T., Keller, P. E., Varlet, M., & Nozaradan, S. (2020). Neural and behavioral evidence for frequency-selective context effects in rhythm processing in humans. *Cerebral Cortex Communications*, *1*, 1–15. https://doi.org/10.1093/texcom/tgaa037

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*, 20200325. https://doi.org/10.1098/rstb.2020.0325

Lenc, T., Peter, V., Hooper, C., Keller, P. E., Burnham, D., & Nozaradan, S. (2022). Infants show enhanced neural responses to musical meter frequencies beyond low-level features. *Developmental Science*, *26*, e13353.

Ley, A., Vroomen, J., & Formisano, E. (2014). How learning to abstract shapes neural sound representations. *Frontiers in Neuroscience*, *8*, 132. https://doi.org/10.3389/fnins.2014.00132

Li, Q., Liu, G., Wei, D., Liu, Y., Yuan, G., & Wang, G. (2019). Distinct neuronal entrainment to beat and meter: Revealed by simultaneous EEG-fMRI. *NeuroImage*, *194*, 128–135. https://doi.org/10.1016/j.neuroimage.2019.03.039

Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, *52*, 57–72. https://doi.org/10.1016/j.neuropsychologia.2013.10.022

Lochy, A., Jacques, C., Maillard, L., Colnat-Coulbois, S., Rossion, B., & Jonas, J. (2018). Selective visual representation of letters and words in the left ventral occipito-temporal cortex with intracerebral recordings. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, E7595–E7604. https://doi.org/10.1073/pnas.1718987115

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199744374.001.0001

London, J., Polak, R., & Jacoby, N. (2017). Rhythm histograms and musical meter: A corpus study of Malian percussion music. *Psychonomic Bulletin and Review*, *24*, 474–480. https://doi.org/10.3758/s13423-016-1093-7

Loseby, P. N., Piek, J. P., & Barrett, N. C. (2001). The influence of speed and force on bimanual finger tapping patterns. *Human Movement Science*, *20*, 531–547. https://doi.org/10.1016/S0167-9457(01)00066-5

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*, *20*, 1133–1139. https://doi.org/10.3758/s13423-013-0439-7

Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199990825.001.0001

Martens, P. A. (2011). The ambiguous tactus: Tempo, subdivision benefit, and three listener strategies. *Music Perception*, *28*, 433–448. https://doi.org/10.1525/mp.2011.28.5.433

Mattioni, S., Rezk, M., Battal, C., Bottini, R., Cuculiza Mendoza, K. E., Oosterhof, N. N., & Collignon, O. (2020). Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *eLife*, *9*, e50732. https://doi.org/10.7554/eLife.50732

Mattioni, S., Rezk, M., Battal, C., Vadlamudi, J., & Collignon, O. (2022). Impact of blindness onset on the representation of sound categories in occipital and temporal cortices. *eLife*, *11*, e79370. https://doi.org/10.7554/eLife.79370

McDermott, J. H., & Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis. *Neuron*, *71*, 926–940. https://doi.org/10.1016/j.neuron.2011.06.032

McKinney, M. F., & Moelants, D. (2006). Ambiguity in tempo perception: What draws listeners to different metrical levels? *Music Perception*, *24*, 155–166. https://doi.org/10.1525/mp.2006.24.2.155

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., Howard, R. M., Hartshorne, J. K., Jennings, M. V., Simson, J., Bainbridge, C. M., Pinker, S., O'Donnell, T. J., Krasnow, M. M., & Glowacki, L. (2019). Universality and diversity in human song. *Science*, *366*, eaax0868. https://doi.org/10.1126/science.aax0868

Meigen, T., & Bach, M. (2000). On the statistical significance of electrophysiological steady-state responses. *Documenta Ophthalmologica*, *98*, 207–232. https://doi.org/10.1023/A:1002097208337

Merchant, H., & Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *Journal of Neuroscience*, *37*, 4552–4564. https://doi.org/10.1523/JNEUROSCI.0367-17.2017

Merchant, H., & Bartolo, R. (2018). Primate beta oscillations and rhythmic behaviors. *Journal of Neural Transmission*, *125*, 461–470. https://doi.org/10.1007/s00702-017-1716-9

Merchant, H., Bartolo, R., Pérez, O., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2014). Neurophysiology of timing in the hundreds of milliseconds: Multiple layers of neuronal clocks in the medial premotor areas. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (pp. 143–154). Springer. https://doi.org/10.1007/978-1-4939-1782-2_8

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, *7*, 274. https://doi.org/10.3389/fnins.2013.00274

Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2015). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *European Journal of Neuroscience*, *41*, 586–602. https://doi.org/10.1111/ejn.12811

Miller KJ, Sorensen LB, Ojemann JG, Den Nijs M (2009) Power-law scaling in the brain surface electric potential Sporns O, ed. *PLoS Computational Biology* 5:e1000609, https://doi.org/10.1371/journal.pcbi.1000609

Milne, A. J., Dean, R. T., & Bulger, D. (2023). The effects of rhythmic structure on tapping accuracy. *Attention, Perception, & Psychophysics*, *85*, 2673–2699. https://doi.org/10.3758/s13414-023-02778-2

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Moshel, M. L., Robinson, A. K., Carlson, T. A., & Grootswagers, T. (2022). Are you for real? Decoding realistic AI-generated faces from neural activity. *Vision Research*, *199*, 108079. https://doi.org/10.1016/j.visres.2022.108079

Niv, Y. (2021). The primacy of behavioral research for understanding the brain. *Behavioral Neuroscience*, *135*, 601–609. https://doi.org/10.1037/bne0000471

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, *15*, 4. https://doi.org/10.1167/15.6.4

Norcia, A. M., Tyler, C. W., Hamer, R. D., & Wesemann, W. (1989). Measurement of spatial contrast sensitivity with the swept contrast VEP. *Vision Research*, *29*, 627–637. https://doi.org/10.1016/0042-6989(89)90048-5

Norman-Haignere, S., Kanwisher, N. G., Mcdermott Correspondence, J. H., & Mcdermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition article distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*, 1281–1296. https://doi.org/10.1016/j.neuron.2015.11.035

Norman-Haignere, S. V., Long, L. K., Devinsky, O., Doyle, W., Irobunda, I., Merricks, E. M., Feldstein, N. A., McKhann, G. M., Schevon, C. A., Flinker, A., & Mesgarani, N. (2022). Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nature Human Behaviour*, *6*, 455–469. https://doi.org/10.1038/s41562-021-01261-y

Nourski, K. V., & Howard, M. A. (2015). Invasive recordings in the human auditory cortex. *Handbook of Clinical Neurology*, *129*, 225–244. https://doi.org/10.1016/B978-0-444-62630-1.00013-5

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging.

Philosophical Transactions of the Royal Society B: Biological Sciences, *369*, 1–10. https://doi.org/10.1098/rstb.2013.0393

Nozaradan, S., Keller, P. E., Rossion, B., & Mouraux, A. (2017). EEG frequency-tagging and input–output comparison in rhythm perception. *Brain Topography*, *31*, 153–160. https://doi.org/10.1007/s10548-017-0605-8

Nozaradan, S., Mouraux, A., Jonas, J., Colnat-Coulbois, S., Rossion, B., & Maillard, L. (2016). Intracerebral evidence of rhythm transform in the human auditory cortex. *Brain Structure and Function*, *222*, 2389–2404. https://doi.org/10.1007/s00429-016-1348-0

Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. *Scientific Reports*, *6*, 20612. https://doi.org/10.1038/srep20612

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, *31*, 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, *32*, 17572–17581. https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Nozaradan, S., Schönwiesner, M., Keller, P. E., Lenc, T., & Lehmann, A. (2018). Neural bases of rhythmic entrainment in humans: Critical transformation between cortical and lower-level representations of auditory rhythm. *European Journal of Neuroscience*, *47*, 321–332. https://doi.org/10.1111/ejn.13826

Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex*, *95*, 156–168. https://doi.org/10.1016/j.cortex.2017.08.015

Okawa, H., Suefusa, K., & Tanaka, T. (2017). Neural entrainment to auditory imagery of rhythms. *Frontiers in Human Neuroscience*, *11*, 493. https://doi.org/10.3389/fnhum.2017.00493

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464. https://doi.org/10.2307/40285633

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*, 827–830. https://doi.org/10.1016/j.cub.2009.03.038

Perez, O., Kass, R. E., & Merchant, H. (2013). Trial time warping to discriminate stimulus-related from movement-related neural activity. *Journal of Neuroscience Methods*, *212*, 203–210. https://doi.org/10.1016/j.jneumeth.2012.10.019

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*, 1430. https://doi.org/10.1126/science.1110922

Picton, T. W. (2010). *Human auditory evoked potentials*. Plural Publishing.

Pikovsky, A., Kurths, J., Rosenblum, M., & Kurths, J. (2003). *Synchronization: A universal concept in nonlinear sciences*. Cambridge University Press.

Polak, R., & London, J. (2014). Timing and meter in Mande drumming from Mali. Music theory. *Online*, *20*. https://doi.org/10.30535/mto.20.1.1

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., & London, J. (2018). Rhythmic prototypes across cultures: A comparative study of tapping synchronization. *Music Perception*, *36*, 1–23.

Povel, D.-J., & Essens, P. J. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–440. https://doi.org/10.2307/40285311

Povel, D.-J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, *30*, 565–572. https://doi.org/10.3758/BF03202011

Rajendran, V. G., Harper, N. S., Garcia-Lazaro, J. A., Lesica, N. A., & Schnupp, J. W. H. (2017). Midbrain adaptation may set the stage for the perception of musical beat. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20171455. https://doi.org/10.1098/rspb.2017.1455

Rajendran, V. G., Harper, N. S., & Schnupp, J. W. H. (2020). Auditory cortical representation of music favours the perceived beat. *Royal Society Open Science*, *7*, 191194. https://doi.org/10.1098/rsos.191194

Rajendran, V. G., & Schnupp, J. W. H. (2019). Frequency tagging cannot measure neural tracking of beat or meter. *Proceedings of the National Academy of Sciences*, *116*, 2779–2780. https://doi.org/10.1073/pnas.1820020116

Ravignani, A., & Norton, P. (2017). Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, *2*, 4–19. https://doi.org/10.1093/jole/lzx002

Regan, D. (1989). *Evoked potentials and evoked magnetic fields in science and medicine*. Elsevier.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992. https://doi.org/10.3758/BF03206433

Repp, B. H. (2008). Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research*, *72*, 79–98. https://doi.org/10.1007/s00426-006-0067-1

Repp, B. H., & Jendoubi, H. (2009). Flexibility of temporal expectations for triple subdivision of a beat. *Advances in Cognitive Psychology*, *5*, 27–41. https://doi.org/10.2478/v10053-008-0063-7

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, *20*, 403–452. https://doi.org/10.3758/s13423-012-0371-2

Retter, T. L., Rossion, B., & Schiltz, C. (2021). Harmonic amplitude summation for frequency-tagging analysis. *Journal of Cognitive Neuroscience*, *33*, 2372–2393. https://doi.org/10.1162/jocn_a_01763

Rosenblum, M. G., Pikovsky, A. S., & Kurths, J. (2001). Phase synchronization: From theory to data analysis. *Handbook of Biological Physics*, *4*, 279–321. https://doi.org/10.1016/S1383-8121(01)80012-9

Rossion, B., Jacques, C., & Jonas, J. (2023). Intracerebral electrophysiological recordings to understand the neural basis of human face recognition. *Brain Sciences*, *13*, 354. https://doi.org/10.3390/brainsci13020354

Rossion, B., & Retter, T. L. (2020). Face perception. In D. Poeppel, G. R. Mangun, & M. S. Gazzaniga (Eds.), *The cognitive neurosciences* (pp. 129–140). MIT Press. https://doi.org/10.7551/mitpress/11442.003.0017

Rossion, B., Retter, T. L., & Liu-Shuang, J. (2020). Understanding human individuation of unfamiliar faces with oddball fast periodic visual stimulation and electroencephalography. *European Journal of Neuroscience*, *52*, 4283–4344. https://doi.org/10.1111/ejn.14865

Rosso, M., Leman, M., & Moumdjian, L. (2021). Neural entrainment meets behavior: The stability index as a neural outcome measure of auditory-motor coupling. *Frontiers in Human Neuroscience*, *15*, 668918. https://doi.org/10.3389/fnhum.2021.668918

Rosso, M., Moens, B., Leman, M., & Moumdjian, L. (2023). Neural entrainment underpins sensorimotor synchronization to dynamic rhythmic stimuli. *NeuroImage*, *277*, 120226. https://doi.org/10.1016/j.neuroimage.2023.120226

Sankaran, N., Swaminathan, J., Micheyl, C., Kalluri, S., & Carlile, S. (2018). Tracking the dynamic representation of consonants from auditory periphery to cortex. *The Journal of the Acoustical Society of America*, *144*, 2462–2472. https://doi.org/10.1121/1.5065492

Sauvé, S. A., Phillips, E., Schiefelbein, W., Daikoku, H., Hegde, S., & Moore, S. (2023). Anti-colonial strategies in cross-cultural music science research. *Music Perception: an Interdisciplinary Journal*, *40*, 277–292. https://doi.org/10.1525/mp.2023.40.4.277

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, *112*, 8987–8992. https://doi.org/10.1073/pnas.1414495112

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, *19*, 831–836. https://doi.org/10.1016/j.cub.2009.03.061

Schulze, H.-H. (1989). Categorical perception of rhythmic patterns. *Psychological Research*, *51*, 10–15. https://doi.org/10.1007/BF00309270

Sethares, W. A., & Staley, T. W. (2001). Meter and periodicity in musical performance. *Journal of New Music Research*, *30*, 149–158. https://doi.org/10.1076/jnmr.30.2.149.7111

Shatek, S. M., Robinson, A. K., Grootswagers, T., & Carlson, T. A. (2022). Capacity for movement is an organisational principle in object representations. *NeuroImage*, *261*, 119517. https://doi.org/10.1016/j.neuroimage.2022.119517

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, *1*, 1–17. https://doi.org/10.1016/0010-0285(70)90002-2

Sifuentes-Ortega, R., Lenc, T., Nozaradan, S., & Peigneux, P. (2022). Partially preserved processing of musical rhythms in REM but not in NREM sleep. *Cerebral Cortex*, *32*, 1508–1519. https://doi.org/10.1093/cercor/bhab303

Slaney, M. (1998). Auditory toolbox, version 2. *Interval Research Corporation, Technical Report*, *1998–010*, 1–52.

Smith, J. O. (2007). *Mathematics of the discrete Fourier transform (DFT): With audio applications*. BookSurge Publishing.

Strasburger, H. (1987). The analysis of steady state evoked potentials revisited. *Clinical Vision Sciences*, *1*, 245–256.

Su, Y. H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, *76*, 373–382. https://doi.org/10.1007/s00426-011-0346-3

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Zion Golumbic, E. (2017). Neural entrainment to the beat: The "missing-pulse" phenomenon. *Journal of Neuroscience*, *37*, 6331–6341. https://doi.org/10.1523/JNEUROSCI.2500-16.2017

Temperley, D., & Bartlette, C. (2002). Parallelism as a factor in metrical analysis. *An Interdisciplinary Journal*, *20*, 117–149. https://doi.org/10.1525/mp.2002.20.2.117

Tierney, A., & Kraus, N. (2014). Neural entrainment to the rhythmic structure of music. *Journal of Cognitive Neuroscience*, *27*, 400–408. https://doi.org/10.1162/jocn_a_00704

Toiviainen, P., & Carlson, E. (2022). Embodied meter revisited: Entrainment, musical content, and genre in music-induced movement. *Music Perception*, *39*, 249–267. https://doi.org/10.1525/mp.2022.39.3.249

Toiviainen, P., & Eerola, T. (2006). Autocorrelation in meter induction: The role of accent structure. *The Journal of the Acoustical Society of America*, *119*, 1164–1170. https://doi.org/10.1121/1.2146084

Toiviainen, P., & Snyder, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception*, *21*, 43–80. https://doi.org/10.1525/mp.2003.21.1.43

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, *28*, 59–70.

Tzanetakis, G., & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Trans Speech Audio Process*, *10*, 293–302. https://doi.org/10.1109/TSA.2002.800560

Van Der Steen, M. C., Jacoby, N., Fairhurst, M. T., & Keller, P. E. (2015). Sensorimotor synchronization with tempo-changing auditory sequences: Modeling temporal adaptation and anticipation. *Brain Research*, *1626*, 66–87. https://doi.org/10.1016/j.brainres.2015.01.053

Volfart, A., Jonas, J., Maillard, L., Colnat-Coulbois, S., & Rossion, B. (2020). Neurophysiological evidence for crossmodal (face-name) person-identity representation in the human left ventral temporal cortex. *PLoS Biology*, *18*, e3000659. https://doi.org/10.1371/journal.pbio.3000659

Wang, X. (2018). Cortical coding of auditory features. *Annual Review of Neuroscience*, *41*, 527–552. https://doi.org/10.1146/annurev-neuro-072116-031302

Wang, Y., Duan, W., Dong, D., Ding, L., & Lei, X. (2022). A test-retest resting, and cognitive state EEG dataset during multiple subject-driven states. *Scientific Data*, *9*, 566. https://doi.org/10.1038/s41597-022-01607-9

Wardle, S. G., Taubert, J., Teichmann, L., & Baker, C. I. (2020). Rapid and dynamic processing of face pareidolia in the human brain. *Nature Communications*, *11*, 4518. https://doi.org/10.1038/s41467-020-18325-8

Weineck, K., Wen, O. X., & Henry, M. J. (2022). Neural entrainment is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *eLife*, *11*, e75515. https://doi.org/10.7554/eLife.75515

Wen, H., & Liu, Z. (2016). Separating fractal and oscillatory components in the power Spectrum of neurophysiological signal. *Brain Topography*, *29*, 13–26. https://doi.org/10.1007/s10548-015-0448-0

Windsor, W. L. (1993). Dynamic accents and the categorical perception of metre. *Psychology of Music*, *21*, 127–140. https://doi.org/10.1177/030573569302100203

Witek, M. A. G. (2017). Filling in: Syncopation, pleasure and distributed embodiment in groove. *Music Analysis*, *36*, 138–160. https://doi.org/10.1111/musa.12082

Witek, M. A. G., Clarke, E. F., Wallentin, M., Kringelbach, M. L., & Vuust, P. (2014). Syncopation, body-movement and pleasure in groove music. *PLoS ONE*, *9*, e94446. https://doi.org/10.1371/journal.pone.0094446

Wojtczak, M., Mehta, A. H., & Oxenham, A. J. (2017). Rhythm judgments reveal a frequency asymmetry in the perception and neural coding of sound synchrony. *Proceedings of the National Academy of Sciences*, *114*, 1201–1206. https://doi.org/10.1073/pnas.1615669114

Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). Interpretations of frequency domain analyses of neural entrainment: Periodicity, fundamental frequency, and harmonics. *Frontiers in Human Neuroscience*, *10*, 274. https://doi.org/10.3389/fnhum.2016.00274

Zuk, N. J., Carney, L. H., & Lalor, E. C. (2018). Preferred tempo and low-audio-frequency bias emerge from simulated sub-cortical processing of sounds with a musical beat. *Frontiers in Neuroscience*, *12*, 349. https://doi.org/10.3389/fnins.2018.00349

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lenc, T., Lenoir, C., Keller, P. E., Polak, R., Mulders, D., & Nozaradan, S. (2025). Measuring self-similarity in empirical signals to understand musical beat perception. *European Journal of Neuroscience*, *61*(2), e16637. https://doi.org/10.1111/ejn.16637