Hebbian learning with elasticity explains how the spontaneous motor tempo affects music performance synchronization

Iran R. Roman, Adrian S. Roman, Ji Chul Kim, Edward W. Large

1 Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, California, United States of America, 2 Department of Mathematics, University of California Davis, Davis, California, United States of America, 3 Department of Psychological Sciences, University of Connecticut, Storrs, Connecticut, United States of America, 4 Department of Physics, University of Connecticut, Storrs, Connecticut, United States of America

* iran@ccrma.stanford.edu

Abstract

A musician’s spontaneous rate of movement, called spontaneous motor tempo (SMT), can be measured while spontaneously playing a simple melody. Data shows that the SMT influences the musician’s tempo and synchronization. In this study we present a model that captures these phenomena. We review the results from three previously-published studies: solo musical performance with a pacing metronome tempo that is different from the SMT, solo musical performance without a metronome at a tempo that is faster or slower than the SMT, and duet musical performance between musicians with matching or mismatching SMTs. These studies showed, respectively, that the asynchrony between the pacing metronome and the musician’s tempo grew as a function of the difference between the metronome tempo and the musician’s SMT, musicians drifted away from the initial tempo toward the SMT, and duet musical performance between musicians with matching or mismatching SMTs. We hypothesize that the SMT constantly acts as a pulling force affecting musical actions at a tempo different from a musician’s SMT. To test our hypothesis, we developed a model consisting of a non-linear oscillator with Hebbian tempo learning and a pulling force to the model’s spontaneous frequency. While the model’s spontaneous frequency emulates the SMT, elastic Hebbian learning allows for frequency learning to match a stimulus’ frequency. To test our hypothesis, we first fit model parameters to match the data in the first of the three studies and asked whether this same model would explain the data the remaining two studies without further tuning. Results showed that the model’s dynamics allowed it to explain all three experiments with the same set of parameters. Our theory offers a dynamical-systems explanation of how an individual’s SMT affects synchronization in realistic music performance settings, and the model also enables predictions about performance settings not yet tested.
Individuals can keep a musical tempo on their own or timed by another individual or a metronome. Experiments show that individuals show a specific spontaneous rate of periodic action, for example walking, blinking, or singing. Moreover, in a simple metronome synchronization task, an individual’s spontaneous rate determines that the individual will tend to anticipate a metronome that is slower, and lag a metronome that is faster. Researchers have hypothesized the mechanisms explaining how spontaneous rates affect synchronization, but no hypothesis can account for all observations yet. Our hypothesis is that individuals rely on adaptive frequency learning during synchronization tasks to adapt the rate of their movements and match another individual’s actions or metronome tempo. Adaptive frequency learning also explains why an individual’s spontaneous rate persists after carrying out a musical synchronization task. We define a new model with adaptive frequency learning and use it to simulate existing empirical data. Not only can our model explain the empirical data, but it can also make testable predictions. Our results support the theory that the brain’s endogenous rhythms give rise to spontaneous rates of movement, and that learning dynamics interact with such brain rhythms to allow for flexible synchronization.

Introduction

Humans can effortlessly sing or walk showing a spontaneous singing tempo or walking pace. In everyday life, however, they also synchronize with external signals, like singing with a pre-recorded song or marching in a parade with other individuals. Humans can adapt the frequency of their actions to align with a common tempo kept by others. This requires perception-action coordination (PAC) [1], involving communication between the brain’s sensory and motor areas [2]. Nonetheless, how an individual’s spontaneous rate of action affects PAC is still an open question [3].

Before reviewing previous research, we wish to clarify some terminology. In the existing literature, stimulus timing (i.e. a metronome) is described in terms of inter-onset-intervals (IOIs) measured in (milli)seconds. Similarly, “musical tempo” is sometimes described as the time between beats, or the inter-beat interval (IBI). In contrast, in the music literature the “tempo” is a rate (bpm: beats per minute), or frequency, not a time period. The relationship between a frequency $f$ (Hz or cycles per second) and a time period $T$ (seconds) is $f = \frac{1}{T}$.

Because we will present a model with an explicit frequency term, when we talk about the SMT (spontaneous motor tempo) and the SPR (spontaneous performance rate) we will use units of Hz. On the other hand, when we talk about the SMP (spontaneous motor period) and the spontaneous performance period (SPP) we will use units of milliseconds. We hope that by acknowledging these historical misnomers we can avoid confusion when trying to understand our model. Table 1 shows a summary of the terms with corresponding abbreviations and units.

The SMT can be measured by asking an individual to spontaneously tap or play a simple melody. It tends to be faster in early childhood compared to adulthood [4] and slower in adult musicians (\(\sim 2.5\text{ Hz on average}\)) compared to non-musicians (\(\sim 3.3\text{ Hz on average}\)) [5, 6]. The mean asynchrony (MA) can be measured in a synchronization task with a metronome where the average time difference between an individual’s actions and the metronome is calculated [7, 8]. Recent studies have investigated how the SMT and the MA are related. In one study, musicians performed a melody while synchronizing with a metronome faster or slower...
than their SMT. Results showed that the MA grew as a function of the difference between the metronome tempo and the musician’s SMT. The MA tended to be positive (musicians lagged) when metronome was faster than the SMT, and negative (musicians anticipated) when the metronome was slower than the SMT. Another study looked at musicians performing a melody without a metronome, starting at a tempo different to the SMT. Results showed a tendency to slowly drift back to the SMT [3], an observation also reported in other studies [4, 9]. A different study looked at duet musical performances where pairs of musicians had matching or mismatching SMTs. Matching duets showed smaller MAs compared to mismatching duets. Each musician’s SMT was remeasured after the performance, revealing that the duet synchronization task did not alter each musician’s SMT [10].

These studies highlight the relationship between the SMT, and the MA. However, currently there only exist models that independently explain the underlying mechanisms of SMT and the MA. The SMT has been hypothesized to originate from central pattern generators in the nervous system [5, 11–13] and motor resonance governed by anatomical properties like body and limb length [14]. The MA has been explained via delayed recurrent feedback in central-peripheral communication between the auditory and motor systems [15–17] and under- or over-estimation of IOI lengths [18]. A potential way to jointly explain the SMT and the MA could be with non-linear oscillators. While the SMT is equivalent to an oscillator’s natural frequency [4, 19, 20], the MA is analogous to the phase difference between an oscillator and a sinusoidal stimulus once they have reached steady-state phase-locked synchronization. This difference shrinks as the oscillator’s natural frequency and the stimulus frequency become closer [21, 22]. Moreover, if the stimulus ceases the oscillator will return to its natural frequency [21, 22].

This study presents a model able to explain behavioral data that relates the SMT and MA. We use the normal form of an Andronov-Hopf bifurcation [23] as described by Large et al. [24]:

$$\tau \frac{dz}{dt} = z(\alpha + i2\pi + \beta|z|^2) + x(t),$$

(1)

where $z$ is the complex-valued state of a non-linear oscillator, $\alpha$ and $\beta$ control its dynamics, $\tau = \frac{1}{f}$ determines its rate ($f$ is frequency in Hz), and $x(t)$ is an external stimulus (Large et al. also included a $\beta_2$ parameter that controls higher-order non-linear activity but we do not use it in the present study [24]). Our model hypothesizes that the SMT originates from central pattern generators [11, 25] and is simulated by Eq 1, which we use in one of its parameter regimes (the supercritical branch of Hopf) due to its limit-cycle properties [24]. However, since Eq 1 can only synchronize with stimulus frequencies close to $f$, its ability to explain both the SMT and the MA is limited [21, 22]. This is clearly distinct from the synchronization abilities of humans, who can synchronize with stimuli tempi that are relatively far away from their SMT [25]. To address this issue we equip Eq 1 with a frequency learning rule.

| Term                        | abbreviation | Units          |
|-----------------------------|--------------|----------------|
| frequency                   | $f$          | Hertz (Hz)     |
| Musical tempo               | N/A          | beats per minute (BPM) |
| Inter-onter interval        | IOI          | (milli)seconds |
| Inter-beat interval         | IBI          | (milli)seconds |
| Spontaneous motor period    | SMP          | (milli)seconds |
| Spontaneous performance period | SPP       | (milli)seconds |
| Spontaneous motor tempo     | SMT          | Hz             |
| Spontaneous performance rate | SPR          | Hz             |

https://doi.org/10.1371/journal.pcbi.1011154.t001
The SMT could be described as an attractor state that pulls human adaptive synchronization to a rate of activity for optimal energy use [4, 25, 26]. Righetti et al. described *dynamic Hebbian learning*, allowing an oscillator to adapt its frequency and match a frequency component present in a stimulus [27]. Consistent with Righetti et al. [28] we call this mechanism “Hebbian learning” due to its similarities with correlation-based adaptation in neural networks, but we acknowledge that the timescale of *dynamic Hebbian learning* is much faster than the slow and long-term changes usually associated with “Hebbian learning”. However, unlike humans [25], Righetti’s oscillators do not return to the initial spontaneous frequency after stimulation ceases. This issue has been addressed by other modeling studies that added elasticity to frequency and phase adaptation models [29–31]. For the specific case of the Andronov-Hopf bifurcation, Lambert et al. added linear elasticity [31] to Righetti’s equation, resulting in adaptive frequency learning with a constant pull to a rate of activity $\omega_0$:

$$\dot{\omega} = -\frac{\lambda_1}{|z|} x(t) \sin(\angle z) - \frac{\lambda_2}{|z|} \left( \omega - \frac{\omega_0}{\omega_0} \right),$$  \hspace{1cm} (2)

where $\dot{\omega} = 2\pi \dot{f}$ (in radians per second), $\omega_0$ is a fixed “spontaneous frequency” that can simulate the SMT, $\lambda_1$ is the frequency learning rate, and $\lambda_2$ is the elasticity strength pulling $\omega$ to $\omega_0$. Combined, Eqs 1 and 2 result in an oscillator that can learn the stimulus frequency, but with a force constantly pulling to an otherwise “spontaneous frequency” of activity.

We propose ASHLE (Adaptive Synchronization with Hebbian Learning and Elasticity), a model that builds upon Eqs 1 and 2 to explain the underlying neural mechanisms relating the SMT and the MA (ASHLE is mathematically described in the methods section). ASHLE uses two oscillators, “sensory” and “motor”, each simulating the excitatory-inhibitory neural dynamics in sensorimotor networks [13, 24]. While the “sensory” oscillator simulates entrainment of auditory-premotor networks [32–35], the “motor” one simulates entrainment of motor networks and actions in PAC [36–38]. In ASHLE, a stimulus can drive and entrain the “sensory” oscillator, which in turn entrains the “motor” one. ASHLE models the SMT as originating from central pattern generators that act as an attractor state for neural and behavioral activity [12]. Therefore, the “motor” oscillator is constantly pulled to a fixed frequency term that simulates the SMT, resulting in an MA between ASHLE’s “motor” oscillator and the stimulus. Behavioral data shows that humans can synchronize with an ongoing stimulus [25], but that they have a tendency to slowly return to their SMT [3]. ASHLE simulates these two timescales by strongly pulling the “motor” oscillator to the SMT, and weakly pulling the “sensory” oscillator to the instantaneous frequency of the “motor” oscillator. These mechanism are the reason why ASHLE uses two oscillators in the first place, because without it the timescales observed in human data would not be possible (see results for experiment 2 and its parameter analysis in the methods section).

Table 2 presents an overview of the most important ASHLE parameters. ASHLE’s oscillators share $\alpha$ and $\beta$ to show spontaneous limit-cycle activity. They also share the frequency learning rate $\lambda_1$, which controls the *dynamic Hebbian learning* allowing the “sensory” and “motor” oscillators to entrain with a stimulus. $f_m$, the dynamic frequency of the “motor” oscillator, is pulled to a spontaneous rate of activity $f_0$ with a $\lambda_2$ strength. In contrast, $f_s$, the dynamic frequency of the “sensory” oscillator, is weakly pulled, with strength $\gamma$, to $f_m$.

In this study we use ASHLE to simulate the dynamics of the MA and the SMT. Fig 1 describes the three previously-published behavioral tasks that we simulate [3, 10, 25]. If ASHLE can explain these dynamics, it would be the first dynamical systems model that systematically captures the relationship between the SMT and the MA as observed in behavioral data. Our results also include simulations that yield predictions that can be tested empirically in future
behavioral studies. It is worth noting that ASHLE only simulates the musical beat in a musical performance and not any other spectral features like pitch, harmony or melody content.

**Experiment 1: Solo music performance with a metronome tempo different than the SMT**

**Results**

We simulated the solo task by Scheurich et al. [25] consisting of performance of a simple melody paced by a metronome (Fig 1A). Their experiment had four different experimental conditions: metronome period 30% shorter, 15% shorter, 15% longer, and 30% longer than the musician’s SMP. Fig 2A shows the asynchrony (specifically the “mean adjusted asynchrony”, see methods section for details) between musician and metronome, which was positive (negative) when the metronome period was shorter (longer) than the musician’s SMP. The asynchrony grows as a function of the difference between SMP and metronome period.

A complex-valued sinusoidal stimulus \( x = \exp(i2\pi f_0 t) \) simulated the metronome (\( f_0 \) is the stimulus frequency in Hz). We hypothesized that ASHLE’s frequency learning, controlled by \( \lambda_1 \), will allow its “sensory” and “motor” oscillators to synchronize with an arbitrary stimulus period. We also expected to see an asynchrony between ASHLE’s “motor” oscillator and the stimulus due to \( \lambda_2 \) pulling the “motor” oscillator to \( f_0 \). We simulated 20 different ASHLE models, each with a unique \( f_0 \) value that matched a musician’s SMT, as measured by Scheurich et al. [25]. Fig 2B shows that ASHLE can explain the behavioral data, and shaded bars show ASHLE’s prediction for the same group of musicians performing with metronome periods 45% shorter or longer than their individual SMP. Fig 2C also makes behavioral predictions, by breaking down how different ASHLE models with different \( f_0 \)s synchronize with a metronome period that is 45%, 30%, and 15% shorter or longer than the period of \( f_0 \).

**Discussion**

Fig 2B showed that our model captures the synchronization dynamics observed in Fig 2A. Specifically, when synchronizing with stimuli slower than \( f_0 \), ASHLE showed a negative MA, and a positive MA when synchronizing with stimuli faster than \( f_0 \). This is explained by the asymmetry of entrained \( f_s \) and \( f_m \) around \( f_0 \), and the elastic pull of \( f_m \) to \( f_0 \) by \( \lambda_2 \). In other words, while ASHLE can entrain with the stimulus, the MA results from the pull that makes \( f_m \) always be “shy” from perfectly matching \( f_s \), the stimulus frequency. It is interesting to note that a closer look at Fig 2A reveals mean adjusted asynchrony magnitudes that are slightly asymmetric between faster and slower stimuli, which implies frequency scaling of PAC. ASHLE has explicit mechanisms that allow it to explain this asymmetry. First, ASHLE has frequency-scaling (see Eqs 3b and 3d) previously used on Andronov-Hopf oscillator models [21, 24]. Second, ASHLE’s elastic pull is an exponential function (see Eq 3d). This means that when synchronizing with faster stimuli the pull to its spontaneous frequency is exponentially larger compared
Fig 1. Illustration of the musical tasks and corresponding simulation experiments. (A) The task simulated in experiment 1, in which a musician plays a simple melody with a metronome (top). Illustration of our simulation, in which ASHLE synchronizes with a sinusoidal stimulus (bottom). (B) The task simulated in experiment 2, in which a musician plays a simple melody, without a metronome (top). This specific example shows a performance tempo that periodically became slower due to the musician’s tendency to return to the SMT. Illustration of our simulation, in which ASHLE oscillates without a sinusoidal stimulus and returns to its $f_0$ (bottom). (C) The task simulated in experiment 3, in which pairs of musicians played a simple melody together after hearing four pacing metronome clicks (top). Illustration of our simulation, in which two ASHLE models synchronize with four cycles of a pacing sinusoidal stimulus (greyed-out blue and red lines), and then stimulate each other without the sinusoidal stimulus (solid blue and red lines) (bottom).

https://doi.org/10.1371/journal.pcbi.1011154.g001
to synchronization with slower stimuli. Together, these two mechanisms cause higher $f_s$ and $f_m$ values to amplify the learning rate ($\lambda_1$) and the elastic force ($\lambda_2$), and lower $f_s$ and $f_m$ values to shrink them.

ASHLE also made testable predictions shown in Fig 2B and 2C. While Fig 2B contains predictions if the same group of musicians were to carry out the task synchronizing with a metronome period 45% shorter or longer than the SMP, Fig 2C contains predictions at the individual musician level, simulating the mean adjusted asynchronies that musicians with different SMP would produce when performing a melody with various metronome tempi faster or slower than their SMT. These predictions can be empirically tested to further validate ASHLE’s behavior or better tune its parameters.

**Experiment 2: Unpaced solo performance with a starting tempo different than the SMT**

**Results**

Experiment 1 showed that ASHLE can explain the MA as captured by Scheurich et al. [25]. In this second experiment, we used ASHLE to simulate the data by Zamm et al. [3], who studied what happens when musicians perform a simple melody unpaced, but starting at a tempo different than the SMT (Fig 1B). They tested four experimental conditions: starting performance tempo fast, faster, slow, and slower than the SMT. Fig 3A shows their results, which for each
condition is the average of the slope (specifically the “mean adjusted slope”, see methods section for details) across musicians. Their data shows that when musicians started at a tempo faster than the SMT the mean adjusted slope was positive (musicians slowed down; consecutive IBIs became longer), and when musicians started at a tempo slower than the SMT the mean adjusted slope was negative (musicians sped up; consecutive IBIs became smaller). In general, musicians showed a tendency to return to their SMT [3].

We used the same set of ASHLE parameter values as in experiment 1. The only differences were that stimulus was nullified, ASHLE’s $f_0$ was dictated by a different set of human SMTs, and initial conditions for $f_s(0) = f_m(0)$ matched human data by Zamm et al. [3]. We hypothesized that the “motor” oscillator will show a tendency to return to ASHLE’s $f_0$ as a result of two mechanisms: $\lambda_2$ pulling to $f_m$ to $f_0$, and $\gamma$ pulling $f_s$ to $f_m$. Fig 3B shows that ASHLE can explain the human data observed in Fig 3A. Additionally, this experiment also yielded predictions of human behavior not yet tested shown in Fig 3C.

**Discussion**

Our model was also able to simulate the human data by Zamm et al. [3] where musicians performed a melody without a metronome. ASHLE’s approximation of the data is not perfect but, keeping in mind that the parameters $\lambda_1$ and $\lambda_2$ were optimized with the data in experiment 1, it is close. Hence, results also show the generalization of ASHLE across different datasets and tasks. Since there was no stimulus in this second experiment ($F = 0$), $\gamma = 0.02$ was optimized to control how quickly ASHLE returns to its $f_0$.

Experiment 2 highlights potential mechanisms that explain how the SMT influences tempo maintenance by a musician during a solo performance. The small $\gamma$ value is explained by the behavioral data in Fig 3A, showing that musicians tended to very slowly return to their SMT. In other words, $\gamma$ represents a musician’s slow tendency to return to their SMT in the absence of stimulation. We predict that non-musicians will show a faster tendency to return to their SMT since musical training prepares musicians to maintain an arbitrary tempo throughout a
musical performance [39, 40]. Future experiments could control ASHLE’s $\gamma$ to explain non-musician data.

ASHLE also made testable predictions shown in Fig 3C at the individual musician level when performing a melody starting at various tempi that are faster or slower than the SMP. These predictions can be empirically tested to further validate ASHLE’s dynamics.

**Experiment 3: Duet musical performance between musicians with matching or mismatching SMTs**

**Results**

Experiments 1 and 2 showed that the same ASHLE model can simulate two tasks carried out by solo musicians. In this third experiment, we used ASHLE to simulate another task by Zamm et al. [10] showing how musician duets perform a simple melody four consecutive times (Fig 1C). Musician duets were separated into two experimental groups: pairs with matching SMTs and pairs with mismatching SMTs. Fig 4A shows their results, which consisted of the mean absolute asynchrony between the beats of the two performing musicians in each experimental group, separately for each of the four melody repetitions. Their results show that the mean absolute asynchrony was smaller between musician duets with matching SMTs.

In this third experiment we also used the same ASHLE parameters as in experiments 1 and 2. However, pairs of ASHLE models are weakly coupled serving as input to each other. We also added Gaussian noise to the “motor” oscillator to better match the magnitudes and variances in absolute asynchrony observed in the behavioral data (see model optimization in the methods section). We hypothesized that pairs of ASHLEs will be able to synchronize due to frequency learning. However, the pull of $f_m$ to $f_0$ will result in asynchrony, which we expect to be smaller between ASHLE pairs with similar $f_0$. We simulated this task using 20 pairs of ASHLE models with $f_0$ that matched the 40 musician SMTs measured by Zamm et al. [10]. Fig 4B shows the mean absolute asynchrony observed in our simulations for the same

![Fig 4. Simulation of the mean absolute asynchrony between two musicians with matching or mismatching SMTs during duet musical performance.](https://doi.org/10.1371/journal.pcbi.1011154.g004)
experimental groups and melody repetitions tested by Zamm et al. [10]. Our results are similar to the human data. Additionally, Fig 4C shows musician data predictions by simulating ASHLE duets where one’s \( f_0 \) has a period that is 220ms shorter, 110ms shorter, 10ms shorter, 10ms longer, 110ms longer, and 220ms longer than the other. Results in Fig 4C systematically describe how the mean absolute asynchrony between musician pairs grows as a function of the difference between their SMTs.

Discussion

ASHLE was also able to simulate duet performances (Fig 4A and 4B). To capture the data of experiment 3, it was necessary to add noise to ASHLE’s motor oscillator (see model optimization in the methods section). Without it, ASHLE showed the same qualitative pattern of results, but with a larger difference in magnitude between experimental groups (see Fig 5). When noise was added, ASHLE showed a mean absolute asynchrony around 15ms and 20ms for the duets with matching and mismatching SMTs, respectively. Moreover, the observed standard error in our simulations is similar to the one observed in the human data. This makes sense as the musical performance task has multiple sources of variability, including: two performing musicians receiving feedback from each other, each with variable behavior that will influence the resulting absolute asynchrony between the two.

![Graph showing mean absolute asynchrony between two musicians with matching or mismatching SMTs during duet musical performance.](https://doi.org/10.1371/journal.pcbi.1011154.g005)
To simulate this musical duet task, we also needed smaller connection strength between pairs of ASHLE models (see model optimization in the methods section). This difference is due to the nature of the stimulus. Sinusoidal stimulation of a Hopf oscillator yields stability regimes with phase-locking that widen as a function of stronger forcing [21]. In contrast, two interacting dynamical systems benefit from weaker coupling to exploit emergent resonance higher-order terms (only up to quadratic in ASHLE’s case) [21–23].

ASHLE also made testable predictions about how musician duets would perform in new experimental manipulations. Fig 4C shows that the mean absolute asynchrony between the two musicians will generally vary as a function of the difference between the musician’s SMTs, and also as a function of the specific SMT value in either of the two musicians. These predictions can also be empirically tested to further validate ASHLE.

General discussion

We have presented ASHLE, a model able to simulate human synchronization dynamics. ASHLE can explain how a musician’s SMT results in asynchronies with a pacing metronome during a simple melody performance task (Fig 2). It can also explain the rate of tempo change when a solo musician performs a simple melody without a metronome (Fig 3). Two ASHLE models can also be coupled to explain the asynchrony between two musicians performing a melody as a duet (Fig 4). ASHLE’s adaptive frequency learning mechanism allows it to generalize across these different tasks while allowing for testable predictions of human data not yet collected (Figs 2B, 2C, 3C and 4C).

ASHLE’s elastic frequency learning captures diverse synchronization dynamics

Experiments 1 and 3 showed that the asynchrony between ASHLE and stimulus is the result of the pull of $f_m$ to $f_0$. Experiment 2 showed that, in the absence of a stimulus, ASHLE can start oscillating at a rate higher or slower than its $f_0$, but that the same elastic pull progressively causes it to return to $f_0$ via the slow tendency of $f$ to converge with $f_m$. ASHLE’s behavior is consistent with theoretical accounts highlighting that a dynamical system’s natural frequency (i.e. ASHLE’s $f_0$) is the optimal state for synchronization, even if the system can synchronize at other frequencies [25, 41–43].

An interesting question is where in the brain these mechanisms of adaptive frequency learning and elasticity occur. ASHLE is a working hypothesis of neuroscientific mechanisms underlying PAC. There exist theories that explain the human SMT as originating from central pattern generators [11–13, 25] and ASHLE’s $f_0$ is a parameter and attractor state consistent with this theory. Additionally, research indicates that cortical and subcortical sensorimotor networks synchronize neural activity that reflects the rhythms of a periodic stimulus and PAC [34, 36, 38, 44]. ASHLE’s oscillators show sustained oscillatory activity to simulate entrainment in sensorimotor brain areas when processing a periodic stimulus of an arbitrary frequency [33–35, 45]. Moreover, ASHLE’s “sensory” and “motor” oscillators show activity synchronized with a stimulus, thus simulating neural entrainment to the musical beat [36]. We also consider the peaks of the “motor” oscillator as an indicator of entrained peripheral effectors (i.e., a finger playing a piano) during a musical performance task. In summary, stimuli entrain ASHLE’s “sensory” oscillator that simulates sensory and premotor neural dynamics. This in turn entrains the “motor” oscillator that simulates motor network dynamics and the transformation into motor commands.

While no consensus exists yet about the interplay of the SMT in PAC, ASHLE is a working model of these hypotheses with mechanisms that explain how a constant pull to the rate of
activity of a central pattern generator affects adaptive PAC. Similar models with period and phase correction mechanisms have also been characterized for their flexible and multi-stable phase-locking dynamics [29, 30, 46, 47], and together with ASHLE support our hypothesis of the underlying elastic frequency learning dynamics of PAC.

**ASHLE’s relation to models of the negative mean asynchrony**

Our study is part of a larger effort to characterize human PAC dynamics. The mean asynchrony has been of particular interest for existing models of PAC [15, 16, 47, 48]. It has been shown to be negative for stimulus with IOIs greater than 300ms but smaller than 2000ms [49], growing as a function of IOI duration. This phenomenon is known as the negative mean asynchrony (NMA). For IOIs smaller than 300ms, the MA has been reported to be absent or slightly positive [48, 50], indicating that 300ms is the transition point for MA to go from negative to positive values. Existing modeling work has largely explained the NMA independent of the SMT, with models proposing that the NMA is the result of adaptive synchronization by means of phase and period correction rules at phenomenological [51] and neuromechanistic levels [47, 52]. Other models follow the strong anticipation hypothesis [15], explaining that the NMA results from delayed neural communication in the sensorimotor system [16].

While ASHLE showed both positive and negative asynchronies, its parameters were optimized to simulate the “mean adjusted asynchrony” reported in the behavioral data of Scheurich et al. [25]. Therefore, what seems to be an NMA in Fig 2A (negative-valued bars) is in reality something else and ASHLE cannot be considered to be an NMA model. A future direction could be to add delayed-feedback to ASHLE’s oscillators and frequency adaptation rules to allow it to explain the NMA while also accounting for the SMT. We also invite human researchers to collect and report both the SMT and the NMA of individual participants in future behavioral studies.

**ASHLE’s potential to simulate non-musician data and its current lack of variability**

Behavioral data shows that musical expertise affects the MA, with musicians showing overall smaller MAs compared to non-musicians [25, 53]. While we optimized ASHLE to explain musician data, its parameters can be controlled to account for non-musician behavior. First, we determined ASHLE’s spontaneous frequency using individual musicians’ SMTs. To simulate non-musician data one would only need to change these values to match a non-musician population. Larger MAs by non-musicians [25, 53] also imply a stronger pull to their SMT, which would translate into ASHLE’s parameters $\lambda_2$ and $\gamma$ being larger than the ones we found.

While ASHLE was able to simulate the musician data in experiments 1 and 2, it does so by reaching a steady-state with little or zero variability (see the variance of Figs 2B and 3B). This was a limitation to perfectly simulate the behavioral results since standard errors were overall smaller in ASHLE’s simulations compared to the behavioral data. In empirical studies of PAC by either musicians or non-musicians there exist two sources of variability: (1) variability within each participant’s behavior across time and (2) variability between the behavior of different participant. In its current form ASHLE only allowed us to simulate the first one, using different ASHLE models with different spontaneous frequency values. Future research could add gaussian noise to ASHLE’s activity and characterize its ability to better approximate the variance associated with human data. Our study includes an early attempt of this by adding noise to ASHLE’s “motor” oscillator in experiment 3.
ASHLE’s relation to gradient frequency neural network models

ASHLE builds upon the work by Large et al. [24] and Lambert et al. [31], who proposed Gradient Frequency Neural Networks (GFNNs). Lambert et al. [31] used a bank of oscillators but, in contrast with Large et al. [24], each oscillator had the ability to adapt its frequency to resonate with a dynamic stimulus (musical rhythms in their case). Lambert et al. also used an elasticity term to pull each oscillator to a fixed central frequency and ensure that oscillators did not overlap in frequency with each other. This allowed them to synchronize their network similarly to Large et al. [24] but using significantly less oscillators.

ASHLE is different in many ways. First, its elastic frequency learning is used to simulate synchronization tasks where asynchronies were observed and explained as a function of the SMT. Second, ASHLE works at the level of synchronization with a musical beat (most music has a beat around \( \approx 1.5 \text{ Hz} \)), a specificity that allows ASHLE to directly process a stimulus using its “sensory” oscillator. Third, ASHLE’s rates of activity in its “sensory” and “motor” oscillators are centered around the SMT (\( \approx 2.5 \text{ Hz} \)) and overlap with the delta band of neural oscillation, which has been hypothesized to be critical for the processing of musical rhythms [54]. Therefore, ASHLE is a minimal model that is specific to simulate behavioral. Future research could look into making ASHLE’s more general. We hypothesize, for example, that it will be possible to use a network of oscillators with elastic frequency learning to extract the beat information directly from a “raw” musical signal. If such oscillators are centered around integer ratios of the SMT, the overall oscillatory activity will reflect synchronization with the stimulus but with a constant pull to the SMT.

Conclusion and outlook

We have presented a versatile model of human behavior during a musical beat synchronization task. ASHLE was able to explain empirical data from three different human studies, and related human behavior with endogenous rhythms like the SMT. Future work could test whether ASHLE’s adaptive frequency learning and synchronization allow it to track semi-periodic signals, such as speech envelopes. Furthermore, ASHLE could be used not just to simulate and make testable predictions of human behavior, but also as a tool with which humans can interact in experimental, musical, and therapy settings. ASHLE advances our understanding about how sensory-motor entrainment to stimulus frequencies interacts with the constraints of endogenous rhythms (i.e. the SMT) hypothesized to originate from central pattern generators.

Methods

The ASHLE model

Eqs 3a, 3b, 3c and 3d show the ASHLE model:

\[
\frac{1}{f_s} \dot{z}_s = z_s(x + i2\pi + \beta|z_s|^2) + x(t) \tag{3a}
\]

\[
\dot{f}_s = f_s \left( \lambda_1 \sin(\phi_x - \phi_s) - \gamma \left( \exp \left( \frac{f_s - f_m}{f_m} \right) - 1 \right) \right) \tag{3b}
\]

\[
\frac{1}{f_m} \dot{z}_m = z_m(x + i2\pi + \beta|z_m|^2) + \exp(i\phi_s) \tag{3c}
\]
\[ \dot{f}_m = f_m \left( \gamma_1 \sin(\phi_s - \phi_m) - \gamma_2 \left( \exp \left( \frac{f_m - f_s}{f_0} \right) - 1 \right) \right) \] (3d)

Eqs 3a and 3c are oscillators like the one in Eq 1. Eqs 3a and 3b have a subscript that stands for “sensory”, while Eqs 3c and 3d have a subscript that stands for “motor”. \( \phi \) stands for the instantaneous phase of the stimulus (\( \phi_s \)), “sensory” (\( \phi_s \)) and “motor” (\( \phi_m \)) oscillators. In all simulations we run in this study \( \alpha = 1 \) and \( \beta = -1 \) so that the intrinsic dynamics of Eqs 3a and 3c are a limit-cycle. In the absence of stimulus (i.e., when \( F = 0 \)), Eqs 3a and 3c will show spontaneous and perpetual oscillation [21]. These limit-cycle properties could be achieved with other oscillators. In fact a phase-only (no amplitude term) oscillator could have been used. However, we select Eq 1 due to its neural underpinnings of excitatory-inhibitory oscillation in cortical and subcortical networks [21, 24, 36, 55]. Eqs 3b and 3d are the frequency learning equations for Eqs 3a and 3c, respectively. Eq 3b has a learning term \( \gamma_1 \sin(\phi_s - \phi_m) \) and a slow term \( -\gamma \exp \left( \frac{f_m - f_s}{f_0} \right) - 1 \). The first one learns the frequency of the external stimulus \( x(t) \), while the second is optimized to weakly pull \( f_s \) to \( f_m \). This weak pull is important when \( F = 0 \), allowing Eq 3b to slowly forget an entrained frequency (see parameter analysis in the next subsection). Eq 3d entrains to match the frequency of Eq 3b, but \( \gamma_2 \) is optimized to strongly pull \( f_m \) to \( f_0 \), which is ASHLE’s spontaneous rate of activity.

**Parameter analysis**

Previous studies have investigated how Eq 3a synchronizes with a periodic external stimulus in a phase-locked fashion [21]. We focus on analyzing how adaptive and elastic frequency learning, controlled by parameters \( \gamma_1 \) and \( \gamma_2 \), affects synchronization with a periodic stimulus. The dynamics of \( \gamma \) will be analyzed in the methods subsection of its relevant experiment (experiment 2).

We set ASHLE’s spontaneous frequency \( f_0 = 2.5 \) because the average musician SMP is around 400ms in previous behavioral studies (\( f_0 = \frac{400\text{ms}}{\text{400ms}} = 2.5\text{Hz} \)) [3, 10, 25]. We analyzed how \( \gamma_1 \) and \( \gamma_2 \) affect the phase-locked asynchrony between the stimulus \( x(t) \) and ASHLE’s \( z_m \) when the stimulus period is 45% shorter (\( f_s = \frac{400\text{ms}}{0.55 \cdot 400\text{ms}} \)) and 45% longer (\( f_s = \frac{400\text{ms}}{1.1 \cdot 400\text{ms}} \)) than ASHLE’s \( f_0 \) period. We ran fifty-second-long simulations, each with a unique value for \( \gamma_1 \) and \( \gamma_2 \), and \( x(t) = \exp(i2\pi f_s t) \), where \( f_s \) is the stimulus frequency in Hz and \( t \) is time. Initial conditions were \( z_s(0) = z_m(0) = 0.001 + i0 \), and \( f_s(0) = f_m(0) = f_0 \). \( \gamma = 0 \) because we do not want to study the effect of this small parameter in this first analysis. After each simulation finished, we found the location (in milliseconds) of all peaks in the real part \( x(t) \) and \( z_m \) (Fig 6A). Then, we subtracted the location of the peaks of \( x(t) \) from the location of the peaks of \( z_m \) and averaged the result to obtain the mean asynchrony in milliseconds. If \( x(t) \) and \( z_m \) showed a different number of peaks, that indicated that phase-locked synchronization did not occur between the two. Fig 6B and 6C show the result of this analysis, with black cells incidating that phase-locked synchronization does not occur for certain combinations of \( \gamma_1 \) and \( \gamma_2 \). Not surprisingly, when \( \gamma_1 = 0 \), phase-locked synchronization is never possible. This makes sense since \( \gamma_1 \) is the frequency learning rate that allows ASHLE to adapt its frequency. When \( \gamma_1 = 0 \), phase-locked synchronization can only occur between ASHLE and a stimulus with a frequency that is close to ASHLE’s \( f_0 \) [21]. This analysis also revealed that as the value of \( \gamma_2 \) becomes larger, phase-locked synchronization may not be observed because ASHLE’s \( f_m \) is strongly pulled to the spontaneous frequency \( f_0 \). The size of the asynchrony is modulated by \( \gamma_1 \) and \( \gamma_2 \). As \( \gamma_1 \)
becomes larger, the asynchrony tends to decrease and is sometimes close to zero when $\lambda_2 = 0$. The opposite occurs when $\lambda_2$ grows. This observation reveals that $\lambda_1$ and $\lambda_2$ work in opposing directions. While $\lambda_1$ changes the model’s frequency to match the stimulus frequency, $\lambda_2$ pulls ASHLE’s to $f_0$. Moreover, as ASHLE’s frequency deviates from $f_0$, $\lambda_2$ acts with more strength, while the strength of $\lambda_1$ is not directly affected by the difference between $f_0$ and ASHLE’s frequency. Additionally, when synchronization is observed between ASHLE and the stimulus $x(t)$, the sign of the asynchrony between $z_m$ and $x(t)$ is affected by whether $x(t)$ is faster or slower than ASHLE’s $f_0$, with a tendency to be positive (lagging) and negative (anticipating), respectively.

**Experiment 1: Solo music performance with a metronome tempo different than the SMT**

**Behavioral data for simulation.** In the task by Scheurich et al. [25], 20 musicians individually performed a simple melody while synchronizing with a metronome in four experimental conditions: metronome period 30% shorter, 15% shorter, 15% longer, and 30% longer than their SMP. For each metronome rate, participants performed the melody (“Mary had a little lamb”) four consecutive times (32 beats per repetition, 128 beats total). Experimenters measured the mean adjusted asynchrony between the participant beats and the metronome clicks during the middle two melody repetitions (64 beats total). The mean adjusted asynchrony is the MA observed when the musician performs with a metronome tempo different than the musician’s SMP, minus the MA observed when the musician performs with a metronome that matches the SMP. Scheurich et al. [25] used the mean adjusted asynchrony instead of the MA to assume in their analysis that no MA exists between the musician and the metronome that matches the SMP. Fig 2A shows the behavioral data with the mean adjusted asynchrony (average and standard error) observed across all 20 musicians for each experimental metronome tempo condition. Their results showed that the mean adjusted asynchrony had a tendency to be positive when synchronizing with a metronome faster than the SMP (musician actions...
lagging the metronome), and negative when synchronizing with a metronome slower than the SMP (musicians actions anticipating the metronome). Additionally, the mean adjusted asynchrony grew as a function of the difference between musician SMP and experimental metronome tempo.

**Setup, procedures and measurements.** To obtain the musicians SMPs, we overlaid a grid over figure 4 from the paper by Scheurich et al. [25], which showed each musician’s measured SMP (in their original paper, the authors call the SMP as the “SPR”). This allowed us to precisely digitize each musician’s SMP from the original publication. We simulated 20 measured SMP (in their original paper, the authors call the SMP as the “SPR”). This allowed us to precisely digitize each musician’s SMP from the original publication. We simulated 20 different ASHLE models, all with the same parameter values (see model optimization below) except for $f_0$, which had a period that matched the SMP of a different musician. In all simulations, initial conditions were $z_m(0) = z_m(0) = 0.001 + i0, f_z(0) = f_m(0) = f_0 = \frac{1000\text{ms}}{1.5\times\text{SMPms}}$. Following the procedure in the musician experiment, each ASHLE synchronized during 128 cycles with a pacing complex-valued sinusoidal stimulus $x(t) = \exp(i2\pi f_x t)$ where $f_x$ had a period either 30% shorter ($f_x = \frac{1000\text{ms}}{1.7\times\text{SMPms}}$), 15% shorter ($f_x = \frac{1000\text{ms}}{1.85\times\text{SMPms}}$), 15% longer ($f_x = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), or 30% longer ($f_x = \frac{1000\text{ms}}{1.3\times\text{SMPms}}$) than ASHLE’s $f_0$ period. We also simulated how ASHLE would synchronize with a stimulus period 45% shorter ($f_x = \frac{1000\text{ms}}{1.4\times\text{SMPms}}$) or 45% longer ($f_x = \frac{1000\text{ms}}{1.05\times\text{SMPms}}$) in order to make predictions about how musicians would perform in those additional experimental conditions. After each simulation, we identified the location (in milliseconds) of local maxima in the real part of ASHLE’s $z_m$ and the sinusoidal stimulus. Then, in each simulation we identified the middle 64 peaks for ASHLE and the stimulus and subtracted the location of stimulus peaks from ASHLE peaks to obtain the asynchrony. Averaging these asynchronies in each simulation resulted in the MA for a specific simulation. To obtain the mean adjusted asynchrony, from each MA obtained in the experimental conditions, we subtracted the MA observed when ASHLE synchronized with a sinusoid with a frequency that matched ASHLE’s $f_0$ ($f_x = f_0$), or 30% longer ($f_x = \frac{1000\text{ms}}{1.3\times\text{SMPms}}$) than ASHLE’s $f_0$ period. We also simulated how different ASHLE models with specific $f_0$ periods (linearly spaced between 350ms and 650ms) would carry out this task when synchronizing with a stimulus period that is 45% shorter ($f_x = \frac{1000\text{ms}}{1.4\times\text{SMPms}}$), 30% shorter ($f_x = \frac{1000\text{ms}}{1.5\times\text{SMPms}}$), 15% shorter ($f_x = \frac{1000\text{ms}}{1.85\times\text{SMPms}}$), 15% longer ($f_x = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), 30% longer ($f_x = \frac{1000\text{ms}}{1.7\times\text{SMPms}}$), or 45% longer ($f_x = \frac{1000\text{ms}}{1.4\times\text{SMPms}}$) than ASHLE’s $f_0$ period. Fig 2C shows the results of these simulations, which are predictions of musician data that could be collected to test the accuracy of predictions made by ASHLE.

**Model optimization.** We identified the set of ASHLE parameters that result in asynchronies observed in the data by Scheurich et al. [25]. We ran simulations where ASHLE was stimulated by a different stimulus frequency to measure the MA between an ASHLE model with an $f_0 = 2.5$ (same as our original parameter analysis shown in Fig 6B and 6C) and a sinusoidal stimulus $x(t) = \exp(i2\pi f_x t)$ with six potential period lengths: 45% shorter ($f_x = \frac{1000\text{ms}}{1.4\times\text{SMPms}}$), 30% shorter ($f_x = \frac{1000\text{ms}}{1.5\times\text{SMPms}}$), 15% shorter ($f_x = \frac{1000\text{ms}}{1.65\times\text{SMPms}}$), 15% longer ($f_x = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), 30% longer ($f_x = \frac{1000\text{ms}}{1.75\times\text{SMPms}}$), and 45% longer ($f_x = \frac{1000\text{ms}}{1.5\times\text{SMPms}}$). The parameter analysis in the previous section revealed that values for $\lambda_1$ between 3 and 5, and $\lambda_2$ between 1 and 3 could result in MA values in the range of the mean adjusted asynchrony observed in the study by Scheurich et al. [25] (Fig 6B and 6C). In this analysis we refine our search for $\lambda_1$ and $\lambda_2$ in this range of values. Subplots in Fig 7 show the asynchrony between $z_m$ and $x(t)$ for a different stimulus frequency and a pair of parameter values of $\lambda_1$ and $\lambda_2$. This analysis revealed that the values of $\lambda_1 = 4$ and $\lambda_2 = 2$ result in the range of MA values observed in humans.
Experiment 2: Unpaced solo music performance with a starting tempo different than the SMT

Behavioral data for simulation. In the task by Zamm et al. [3], 24 musicians individually performed a simple melody without listening to a metronome. First they performed the melody at their SMT, and in four other experimental conditions they performed starting at four other spontaneous tempi: fast and slow with respect to SMT, and even faster and slower with respect to the SMT. For each spontaneous initial tempo, musicians performed the melody...
A model explaining the role of endogenous rhythms in musical action

Setup, procedures and measurements. To obtain the musicians’ SMP, we overlaid a grid over Figure 1 from the paper by Zamm et al. [3], which showed each musician’s measured SMP (in their original paper, the authors call the SMP as the “SPR”). We also overlaid a grid over Figure 2 (right top panel) from the same paper by Zamm et al. to obtain each musician’s initial performance tempo values. Using the musicians’ SMP and initial performance tempo values, we simulated 23 different ASHLE models, each with 5 different initial conditions \( f_i(0) = f_m(0) \): (1) matching a musician’s SMP, (2) “fast” or (3) “slow” compared to the SMT, and (4) “faster” and (5) “slower” than the SMT (115 total simulations). We did not simulate the participant with the SMP of 665ms because their “fast” spontaneous tempo was significantly faster than the rest of the participants’ “fast” tempi. ASHLE was not able to show stable activity as a result of this tempo difference. In all simulations in this experiment there was no stimulus (\( F = 0 \)). Other than \( F_0 \) and the initial conditions for \( f_i(0) = f_m(0) \), all simulations shared the same parameter values (see model optimization below). After each simulation, we identified the location (in milliseconds) of local maxima in the real part of ASHLE’s \( z_m \). Then, we measured the difference between consecutive peaks to analyze how IBIs change over the course of a simulation. For each simulation we carried out a linear regression over the IBIs to obtain a slope value. Consistent with the methods in the human experiment to obtain the adjusted slope, the slope of each simulation in the control condition where \( f_i(0) = f_m(0) = f_0 \) was subtracted from the slope obtained in the experimental conditions \( f_i(0) = f_m(0) \neq f_0 \). After accounting for numerical error we observed that ASHLE had slope of virtually zero when \( f_i(0) = f_m(0) = f_0 \). We averaged adjusted slopes across ASHLE models to obtain the plot in Fig 3B. We also simulated how different ASHLE models with specific \( f_0 \) values (with period lengths linearly spaced between 350ms and 650ms) would carry out this task when the initial conditions for \( f_i(0) = f_m(0) \) were a period 45% shorter \( f_i(0) = f_m(0) = \frac{1000}{0.55 \times \text{SMPms}} \), 30% shorter \( f_i(0) = f_m(0) = \frac{1000}{0.85 \times \text{SMPms}} \), 15% longer \( f_i(0) = f_m(0) = \frac{1000}{1.15 \times \text{SMPms}} \), and 45% longer \( f_i(0) = f_m(0) = \frac{1000}{1.45 \times \text{SMPms}} \) than ASHLE’s \( f_0 \) period. Fig 3C shows the results of these simulations, which are predictions of musician data that could be collected to test the accuracy of predictions made by ASHLE.

Model optimization. We identified the parameter \( \gamma \) that results in ASHLE’s return to \( f_0 \) when there is no stimulus present (\( F = 0 \)) and the initial values of \( f_i(0) = f_m(0) \neq f_0 \). To optimize ASHLE we use the same setup as in our original parameter analysis with the exception that initial conditions for \( f_i(0) \) and \( f_m(0) \) were set to one of four different values: period 45% shorter \( f_i(0) = f_m(0) = \frac{1000}{0.55 \times \text{SMPms}} \), 30% shorter \( f_i(0) = f_m(0) = \frac{1000}{0.85 \times \text{SMPms}} \), 15% shorter \( f_i(0) = f_m(0) = \frac{1000}{1.15 \times \text{SMPms}} \), 30% longer \( f_i(0) = f_m(0) = \frac{1000}{1.35 \times \text{SMPms}} \), and 45% longer \( f_i(0) = f_m(0) = \frac{1000}{1.45 \times \text{SMPms}} \) than ASHLE’s \( f_0 \) period. In all simulations there was no external stimulus (i.e., \( F = 0 \)). We analyzed the behavior of ASHLE for \( \gamma \) values of 0.01, 0.02, 0.04, 0.08. In each simulation, ASHLE oscillated for 50
seconds with initial conditions $z_s(0) = z_m(0) = 0.001 + i0$. Next, we found the location (in milliseconds) of the local maxima in the real-part of the oscillatory activity of $z_m$. Then we found the difference between consecutive local maxima to obtain a sequence of IBIs. We calculated the linear regression between consecutive IBIs to obtain the resulting slope. Each line in Fig 8A shows the slopes obtained with different $\gamma$ values and different initial conditions for $f_s(0)$ and $f_m(0)$. This analysis revealed that a value around $\gamma = 0.02$ will match the range of slope values observed in human data.

Finally, to determine if ASHLE really needs two oscillators, we also ran this analysis with an alternative single-oscillator model that consists of ASHLE’s Eqs 3a and 3b where $f_m$ is substituted by $f_0$. Fig 8B shows the resulting slope values, revealing an order of magnitude increase in slope values. Such a single-oscillator model would need separate tuning and switching of the parameter pulling the oscillator to $f_0$ to be able to explain the data in experiments 1 and 2.

ASHLE, in contrast, natively shows two different time-scales depending on whether it is being stimulated or not, allowing it to explain both experiment 1 and 2 without the need for switching or tuning its parameters.

**Experiment 3: Duet musical performance between musicians with matching or mismatching SMTs**

**Behavioral data for simulation.** In another experiment, Zamm et al. [10] measured the SMP of 40 musicians and formed duets in two experimental groups: matching SMPs (<10ms IBI difference) and mismatching SMPs (>110ms IBI difference). There were 10 unique musician duets in each experimental group. Musician pairs were instructed to perform a simple unfamiliar melody together (16 beats in length), repeating the melody four consecutive times.
Because the same melody was repeated four times, they measured each melody repetition separately. Fig 4A shows their behavioral data. The mean absolute asynchrony between duets of each pair of synchronizing musicians throughout the entire performance. Experimenters measured the absolute asynchrony between beats (400ms IOI) that established the common tempo. When pairs of musicians performed the task, they first heard four metronome beats that established the common tempo. Experimenters measured the absolute asynchrony between each pair of synchronizing musicians throughout the entire performance.

Setup, procedures and measurements. To obtain the musicians’ SMPs, we overlaid a grid over Figure 1 from their paper [10], which showed each musician’s measured SMP (in their original paper, the authors call the SMP as the “SPR”). This allowed us to precisely recover each musician’s SMP. We simulated 20 pairs of ASHLE models (10 pairs with similar natural frequencies and 10 pairs with dissimilar natural frequencies) synchronizing during 64 cycles. All simulations shared the same parameter values and initial conditions used in our experiment 1, except for the coupling strength between synchronizing ASHLE models (see model optimization in the next paragraph) and \( f_0 \), which had a period that matched the SMP of a musician in the study. At the beginning of the simulation, two ASHLE models were stimulated by a complex-valued sinusoid \( x(t) = \exp(i2\pi f_z t) \) with a period of 400ms (\( f_z = 2.5 \)). After these four cycles of sinusoidal stimulation, the stimulus stopped and the two ASHLE models stimulated each other with their respective \( z_m \). That is, in each duet simulation, after four cycles of sinusoidal stimulation, the input to the ASHLE No.1 was \( F_z z_m \) and the input to ASHLE No.2 was \( F_z z_m \), where \( F_z \) is the forcing strength between ASHLE models. After each simulation, we identified the location (in milliseconds) of local maxima in the real part of each ASHLE’s \( z_m \). Then, we measured the absolute asynchrony between the two synchronizing ASHLE model’s \( z_m \), obtaining 64 absolute asynchronies for each simulation. We divided these in four subsections (16 absolute asynchronies each) to simulate the four melody repetitions that pairs of musicians carried out, resulting in Fig 4B. We also simulated how different ASHLE models (with \( f_0 \) periods linearly spaced between 350ms and 650ms) would carry out this task when synchronizing for 16 cycles with another ASHLE model with an \( f_0 \) period difference of -220ms, -110ms, -10ms, 10ms, 110ms, and 220ms. Fig 4C shows these simulations, which are predictions of data that could be collected to test the accuracy of predictions made by ASHLE.

Model optimization. There were two kinds of stimulation in this third experiment. During the first four cycles, similar to experiment 1, ASHLE was simulated by a sinusoid \( x(t) = F \exp(i2\pi f_z t) \) with a force \( F = 1 \) and \( f_z = 2.5 \). During the next 64 cycles, two ASHLE models synchronized with each other, so the input to the first ASHLE model is the second ASHLE model’s \( F_z z_m \). We found that using \( F_z = 1 \) resulted in a lack of phase-locked synchronization between the two ASHLE, suggesting that \( F_z = 1 \) is too large and causes unstable dynamics between the two interacting ASHLE models. To improve stability, we reduced the value of \( F_z \) until we observed stable synchronization between all pairs of ASHLE models that we want to simulate, with the optimal \( F_z = 0.01 \). However, we also noted that the mean absolute asynchrony was considerably smaller between pairs of ASHLE models with similar \( f_0 \) compared to the results for musician duets with matching SMTs. We believe that this difference was due to the lack of noise in our model. To improve our simulations we added gaussian noise to ASHLE’s \( z_m \) turning Eq 3c into:

\[
\frac{1}{f_m} \ddot{z}_m = z_m (\chi + i2\pi + \beta|z_m|^2) + \exp(i/z) + N(\mu, \sigma^2),
\]

where \( \mu = 0 \) is the mean and \( \sigma = 10 \) is the standard deviation of a normal distribution. Fig 4B shows the results obtained after we added the noise, which better approximate the behavioral data.
Author Contributions
Conceptualization: Iran R. Roman, Adrian S. Roman, Edward W. Large.
Formal analysis: Iran R. Roman, Ji Chul Kim, Edward W. Large.
Investigation: Iran R. Roman, Adrian S. Roman, Edward W. Large.
Methodology: Iran R. Roman, Ji Chul Kim, Edward W. Large.
Software: Iran R. Roman, Adrian S. Roman.
Supervision: Edward W. Large.
Validation: Iran R. Roman, Ji Chul Kim.
Visualization: Iran R. Roman, Adrian S. Roman, Edward W. Large.
Writing – original draft: Iran R. Roman, Ji Chul Kim, Edward W. Large.
Writing – review & editing: Iran R. Roman, Ji Chul Kim, Edward W. Large.

References
1. Oullier O, Jantzen KJ, Steinberg FL, Kelso JA. Neural substrates of real and imagined sensorimotor coordination. Cerebral Cortex. 2005 Jul 1; 15(7):975–85. https://doi.org/10.1093/cercor/bhh198 PMID: 15563729
2. Riddervik K, R. Neurocognitive mechanisms of perception–action coordination: A review and theoretical integration. Neuroscience & Biobehavioral Reviews. 2014; 46:3–29. https:// doi.org/10.1016/j.neubiorev.2014.05.008 PMID: 24860965
3. Zamm A, Wang Y, Palmer C. Musicians’ natural frequencies of performance display optimal temporal stability. Journal of Biological Rhythms. 2018; 33(4):432–440. https://doi.org/10.1177/0748730418783651 PMID: 29940801
4. McAuley JD, Jones MR, Holub S, Johnston HM, Miller NS. The time of our lives: life span development of timing and event tracking. Journal of Experimental Psychology: General. 2006; 135(3):348. https://doi.org/10.1037/0096-3445.135.3.348 PMID: 16846269
5. Scheurich R, Zamm A, Bogetti C, Palmer C. Spontaneous Production Rates Are Consistent Across Tasks Varying in Motor Complexity. In: Canadian journal of experimental psychology -revue canadienne de Psychologie experimentale. Vol. 70. Canadian psychological assoc 141 laurier Ave west, ste 702, ottawa, ontario . . .; 2016. P. 402–402.
6. Drake C, Penel A, Bigand E. Tapping in time with mechanically and expressively performed music. Music Perception. 2000; 18(1):1–23. https://doi.org/10.2307/40285899
7. Repp BH. Sensorimotor synchronization: a review of the tapping literature. Psychonomic bulletin & review. 2005; 12(6):969–992. https://doi.org/10.3758/BF03206433 PMID: 16615317
8. Repp BH, Su YH. Sensorimotor synchronization: a review of recent research (2006–2012). Psychonomic bulletin & review. 2013; 20(3):403–452. https://doi.org/10.3758/s13423-012-0371-2 PMID: 23397235
9. Yu H, Russell DM, Stenard D. Task-effector asymmetries in a rhythmic continuation task. Journal of Experimental Psychology: Human Perception and Performance. 2003; 29(3):616. PMID: 12848329
10. Zamm A, Wellman C, Palmer C. Endogenous rhythms influence interpersonal synchrony. Journal of Experimental Psychology: Human Perception and Performance. 2016; 42(5):611. PMID: 26820249
11. Latash M. Virtual trajectories, joint stiffness, and changes in the limb natural frequency during single-joint oscillatory movements. Neuroscience. 1992; 49(1):209–220. https://doi.org/10.1016/0306-4522(92)90089-K PMID: 1407547
12. Marder E, Bucher D. Central pattern generators and the control of rhythmic movements. Current biology. 2001 Nov 27; 11(23):R986–96. https://doi.org/10.1016/S0960-9822(01)00581-4 PMID: 11728329
13. Guertin PA. The mammalian central pattern generator for locomotion. Brain research reviews. 2009 Dec 11; 62(1):45–56. https://doi.org/10.1016/j.brainresrev.2009.08.002 PMID: 19720083
14. Goodman L, Riley MA, Mitra S, Turvey M. Advantages of rhythmic movements at resonance: minimal active degrees of freedom, minimal noise, and maximal predictability. Journal of motor behavior. 2000; 32(1):3–8. https://doi.org/10.1080/00222890009601354 PMID: 11082866
15. Stepp N, Turvey MT. On strong anticipation. Cognitive systems research. 2010; 11(2):148–164. https://doi.org/10.1016/j.cogsys.2009.03.003 PMID: 20191086

16. Roman IR, Washburn A, Large EW, Chafe C, Fujioka T. Delayed feedback embedded in perception-action coordination cycles results in anticipation behavior during synchronized rhythmic action: A dynamical systems approach. PLoS computational biology. 2019; 15(10):e1007371. https://doi.org/10.1371/journal.pcbi.1007371 PMID: 31671096

17. Aschersleben G. Temporal control of movements in sensorimotor synchronization. Brain and cognition. 2002; 48(1):66–79. https://doi.org/10.1006/brcg.2001.1304 PMID: 11812033

18. Loehr JD, Palmer C. Subdividing the beat: Auditory and motor contributions to synchronization. Music Perception. 2009; 26(5):415–425. https://doi.org/10.1525/mp.2009.26.5.415

19. Large EW, Fink P, Kelso SJ. Tracking simple and complex sequences. Psychological research. 2002; 66(1):3–17. https://doi.org/10.1006/s004260100009 PMID: 11963276

20. Large EW, Palmer C. Perceiving temporal regularity in music. Cognitive science. 2002; 26(1):1–37. https://doi.org/10.1207/s15516709cog2601_1

21. Kim JC, Large EW. Signal processing in periodically forced gradient frequency neural networks. Frontiers in computational neuroscience. 2015; 9:152. https://doi.org/10.3389/fncom.2015.00152 PMID: 26733858

22. Kim JC, Large EW. Mode locking in periodically forced gradient frequency neural networks. Physical Review E. 2019; 99(2):022421. https://doi.org/10.1103/PhysRevE.99.022421 PMID: 30934299

23. Hoppensteadt FC, Izhikevich EM. Weakly connected neural networks. Springer Science & Business Media; 1997 Jul 10.

24. Large EW, Almonte FV, Velasco MJ. A canonical model for gradient frequency neural networks. Physica D: Nonlinear Phenomena. 2010; 239(12):905–911. https://doi.org/10.1016/j.physd.2009.11.015

25. Scheurich R, Zamm A, Palmer C. Tapping into rate flexibility: musical training facilitates synchronization around spontaneous production rates. Frontiers in psychology. 2018; 9:458. https://doi.org/10.3389/fpsyg.2018.00458 PMID: 29681872

26. SH S, I S. Coupled oscillators and biological synchronization. Scientific American. 1993; 269(6):102–109. https://doi.org/10.1038/scientificamerican1293-102

27. Righetti L, Buchli J, Ijspeert AJ. Dynamic hebbian learning in adaptive frequency oscillators. Physica D: Nonlinear Phenomena. 2006; 216(2):269–281. https://doi.org/10.1016/j.physd.2006.02.009

28. Kempter R, Gerstner W, van Hemmen JL. Hebbian learning and spiking neurons. Phys. Rev. E 1999; 59(99):4498–4514. https://doi.org/10.1103/PhysRevE.59.4498

29. Ermentrout B. An adaptive model for synchrony in the firefly Pteroptyx malaccae. Journal of Mathematical Biology. 1991; 29(6):571–585. https://doi.org/10.1007/BF00164052

30. Savinov M, Swigon D, Ermentrout B. Synchronization and locking in oscillators with flexible periods. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2021 Mar 19; 31(3):033143. https://doi.org/10.1063/5.0021836 PMID: 33810738

31. Lambert AJ, Weyde T, Armstrong N. Adaptive Frequency Neural Networks for Dynamic Pulse and Metre Perception. In: ISMIR. Schloss Dagstuhl LIZI; 2016. p. 60–66.

32. Patel AD, Iversen JR. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Frontiers in systems neuroscience. 2014; 8:57. https://doi.org/10.3389/fnsys.2014.00057 PMID: 24860439

33. Daly I, Hallowell J, Hwang F, Kirke A, Malik A, Roesch E, et al. Changes in music tempo entrain movement related brain activity. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2014. p. 4595–4598.

34. Grahn JA, Rowe JB. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. Journal of Neuroscience. 2009; 29(23):7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009 PMID: 19515922

35. Grahn JA, Rowe JB. Finding and feeling the musical beat: striatal dissociations between detection and prediction of regularity. Cerebral cortex. 2013; 23(4):913–921. https://doi.org/10.1093/cercor/bhs083 PMID: 22499797

36. Large EW, Herrera JA, Velasco MJ. Neural networks for beat perception in musical rhythm. Frontiers in systems neuroscience. 2015; 9:159. https://doi.org/10.3389/fnsys.2015.00159 PMID: 26635549

37. Will U, Berg E. Brain wave synchronization and entrainment to periodic acoustic stimuli. Neuroscience letters. 2007 Aug 31; 424(1):55–60. https://doi.org/10.1016/j.neulet.2007.07.036 PMID: 17709189

38. Fujioka T, Trainor LJ, Large EW, Ross B. Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. Journal of Neuroscience. 2012; 32(5):1791–1802. https://doi.org/10.1523/JNEUROSCI.4107-11.2012 PMID: 22302818
39. Fine P, Bull S. Memory for tactus and musical tempo: The effects of expertise and speed on keeping time. In: Proceedings of the International Symposium on Performance Science; 2009.

40. Schultz BG, Palmer C. The roles of musical expertise and sensory feedback in beat keeping and joint action. Psychological research. 2019; 83(3):419–431. https://doi.org/10.1007/s00426-019-01156-8 PMID: 30805705

41. Von Holst E. On the nature of order in the central nervous system. The Collected Papers of Erich von Holst Vol 1, The Behavioral Physiology of Animal and Man. 1937; p. 133–155.

42. Haken H, Kelso JS, Bunz H. A theoretical model of phase transitions in human hand movements. Biological cybernetics. 1985; 51(5):347–356. https://doi.org/10.1007/BF00336922 PMID: 3978150

43. Kelso J. Relative timing in brain and behavior: Some observations about the generalized motor program and self-organized coordination dynamics. Human Movement Science. 1997; 16(4):453–460. https://doi.org/10.1016/S0167-9457(96)00044-9

44. Chen JL, Penhune VB, Zatorre RJ. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. Journal of cognitive neuroscience. 2008; 20(2):226–239. https://doi.org/10.1162/jocn.2008.20018 PMID: 18275331

45. Large EW, Jones MR. The dynamics of attending: How people track time-varying events. Psychological review. 1999; 106(1):119. https://doi.org/10.1037/0033-295X.106.1.119

46. Loehr JD, Large EW, Palmer C. Temporal coordination and adaptation to rate change in music performance. Journal of Experimental Psychology: Human Perception and Performance. 2011 Aug; 37(4):1292. PMID: 21553990

47. Bose A, Byrne Á, Rinzel J. A neuromechanistic model for rhythmic beat generation. PLoS computational biology. 2019; 15(5):e1006450. https://doi.org/10.1371/journal.pcbi.1006450 PMID: 31071078

48. Repp BH. Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. Journal of motor behavior. 2003; 35(4):355–370. https://doi.org/10.1080/002228903039603156 PMID: 14607773

49. Mates J, Müller U, Radl T, Pöppel E. Temporal integration in sensorimotor synchronization. Journal of cognitive neuroscience. 1994; 6(4):332–340. https://doi.org/10.1162/jocn.1994.6.4.332 PMID: 23961729

50. Wohlschläger A. Synchronization error: An error in time perception. In: Abstracts of the Psychonomic Society. vol. 4; 1999. p. 48.

51. Van Der Steen MC, Keller PE. The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. Frontiers in human neuroscience. 2013; 7:253. https://doi.org/10.3389/fnhum.2013.00253 PMID: 23772211

52. Large EW, Roman I, Kim JC, Cannon J, Pazdera JK, Trainor LJ, Rinzel J, Bose A. Dynamic models for musical rhythm perception and coordination. Frontiers in Computational Neuroscience. 2023 May 17. https://doi.org/10.3389/fncom.2023.1151896 PMID: 37265781

53. Repp BH, Doggett R. Tapping to a very slow beat: a comparison of musicians and nonmusicians. Music Perception. 2007; 24(4):367–376. https://doi.org/10.1525/mp.2007.24.4.367

54. Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppe D. Temporal modulations in speech and music. Neuroscience & Biobehavioral Reviews. 2017 Oct 1; 81:181–7. https://doi.org/10.1016/j.neubiorev.2017.02.011 PMID: 28212857

55. Lerud KD, Almonte FV, Kim JC, Large EW. Mode-locking neurodynamics predict human auditory brainstem responses to musical intervals. Hearing Research. 2014 Feb 1; 308:41–9. https://doi.org/10.1016/j.heares.2013.09.010 PMID: 24091182