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

Iran R. Roman[1*], Adrian S. Roman[2], Ji Chul Kim[3], Edward W. Large[3,4]

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, United States of America
4 Department of Physics, University of Connecticut, Storrs, United States of America

* iran@ccrma.stanford.edu

Abstract

A musician's spontaneous rate of movement, called spontaneous motor tempo (SMT), can be measured when the musician spontaneously plays a simple melody. Data shows that SMT systematically influences tempo and synchronization among musicians. In this study we present a model that captures these phenomena. To develop our model, we review the results from three musical performance settings that have been previously published: (1) solo musical performance with a pacing metronome tempo that is different from the SMT, (2) solo musical performance without a metronome at a tempo that is faster or slower than the SMT, and (3) duet musical performance between musician pairs with matching or mismatching SMTs. In the first, 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. In the second, musicians drifted away from the initial spontaneous tempo toward the SMT. And in the third, the absolute asynchronies between performing musicians were smaller if their SMTs matched compared to when they did not. We hypothesize that, while musicians can perform musical actions at a tempo different from their SMT, the SMT constantly acts as a pulling force. We developed a model to test our hypothesis. The model is an oscillatory dynamical system with Hebbian tempo learning and elasticity that simulates synchronization in music performance. The dynamical system's natural frequency models SMT and Hebbian learning lets the system's frequency adapt to match the stimulus frequency. The pulling force is modeled as an elasticity term that pulls the learned frequency toward the system's natural frequency. To test our hypothesis, we chose model parameters based on the results of the first experiment, and asked whether they could predict the results of the next two experiments using the same set of parameters. The model fit the results of the first experiment closely, and accurately predicted the two additional data sets. Thus, our theory offers a dynamical 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.

Author summary

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 elastic Hebbian frequency learning during synchronization tasks to adapt the rate of their movements and match another individual's actions or metronome tempo. Elastic Hebbian frequency learning also explains why an individual's spontaneous rate persists after carrying out a musical synchronization task. We define a new model with elastic Hebbian 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 a song in the shower or walk down an empty street. In such a solo setting, individuals show a spontaneous singing tempo or walking pace. In everyday life, however, humans often have to synchronize with external signals. For example, singing with a pre-recorded song or marching in a parade with other individuals. In the specific case of musical performances, musicians change the tempo of their actions to match a musical tempo that is kept by other performing musicians. This requires perception-action coordination, and involves coordinated communication between the brain's sensory and motor systems [1]. However, most musical performances do not match a musician's spontaneous tempo. How a musician's spontaneous tempo affects a musical performance is still an open question [2]. To tackle this question, one must first study the spontaneous tempo of periodic action, like finger tapping.

Before reviewing previous research, we wish to clarify some terminology. In the literature on rhythmic synchronization, stimulus timing is historically described in terms of inter-onset-intervals (IOIs), measured in (milli)seconds. For a metronome, timing can be described as the length (in units of seconds) of one metronome cycle, or its period. Similarly musical tempo is sometimes described as the time between beats, or the inter-beat interval (IBI), which again, is the length of a cycle (a period). The problem is that the term "tempo" in music is a frequency (beats per minute, bpm) not a period. When period is measured in seconds, and frequency in cycles per second (ie, Hertz) the relationship is f (Hz) = 1 / T (sec) = 1000 / T (ms). Because we will discuss physics-based models in which frequency is the relevant variable, this distinction is critical. For this reason, when we use SMT (spontaneous motor tempo), we will consider it a frequency, and measurements in terms of milliseconds will be called SMP (spontaneous motor period). Similarly, we will use SPR (spontaneous performance rate) to refer to a frequency, and use spontaneous performance period (SPP) to refer to measurements in units of time. We hope that by acknowledging these historical misnomers we can avoid confusion when trying to understand our physics-based dynamical models. Table 1 shows a summary of the terms, abbreviations, and the correct corresponding units

SMTs vary within individuals, tending to be faster in early childhood compared to

Table 1. Terminology for periodic events

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

adulthood [3], and between individuals, tending to be slower in adult musicians ($\sim$ 2.5 Hz on average) compared to non-musicians ($\sim$ 3.3 Hz on average) [5, 6]. After studying the SMT, one can study synchronization of simple movements, like finger-tapping, with a metronome. The asynchrony between an individual's taps and metronome clicks can be measured, and the mean asynchrony (MA) can be calculated [7, 8]. For both musicians and non-musicians the MA tends to be negative (taps precede the metronome click) when synchronizing with a metronome with an IOI between 300ms and 2000ms [9]. However, the MA becomes smaller as the metronome IOI approaches 300ms and for an IOI smaller than 300ms the MA is absent or even slightly positive [10, 11]. Interestingly, the 300ms IOI, around where the negative MA disappears, coincides with the mean SMT observed in humans ($\sim$ 2.85 Hz), indicating a possible connection between SMT and MA dynamics. However, musical expertise does affect the MA, with musicians showing overall smaller MAs compared to non-musicians [12]. Additionally, the MA is also positive (taps lag the metronome clicks) when synchronizing with a very slow metronome with IOIs greater than 5000ms [13]. Synchronization with a slow metronome is a difficult task, so the positive MA in this case may be the result of human actions reacting to the metronome clicks, rather than a direct relationship between the MA and the SMT [12]. When synchronizing with a metronome with an IOI between 2000ms and 5000ms, human actions form a bimodal distribution (i.e., some taps precede and some lag the metronome clicks) [14]. Together, the SMT and the MA are behavioral measurements we are interested in. While the SMT is a measurement of spontaneous, non-paced action, the MA is a measurement of paced action. Studying both the SMT and the MA lets us better understand how the spontaneous tempo of periodic action affects an individual's synchronization with an external periodic stimulus.

Previous studies have looked at musical performance in solo and group settings to investigate how the SMT and the MA are related to each other. In one study, musicians performed a melody in a solo setting synchronizing with a metronome tempo that was faster or slower than the musician's SMT. Results showed that the MA grew as a function of the difference between the metronome tempo and the musician's SMT. When the metronome was faster than the SMT, the MA had a tendency to be positive (the musician's beat lagged the metronome clicks), and when the metronome was slower than the SMT, the MA had a tendency to be negative (the musician's beat preceded the metronome clicks). The smallest MA was observed when the metronome tempo matched the individual's SMT [15]. In another study, musicians performed a melody in a solo setting without a metronome, starting at spontaneous tempo that was slower or faster than the SMT. Results showed that the musician's tempo had a tendency to drift back to the SMT [2]. This tendency to drift back to the SMT has also been observed in other similar studies [3, 16]. One study has also investigated the relationship between the SMT and the MA in duet musical performances. Musicians were split into two groups: one group consisted of musician pairs with matching SMT (difference in period duration smaller than 10ms) and the other group consisted of musician pairs with

mismatching SMT (difference in period duration greater than 110ms). Next, pairs of musicians were asked to perform a melody with each other. Musician duets with matching SMTs showed a smaller mean absolute asynchrony compared to synchronizing pairs with mismatching SMTs. Additionally, after the duet performance each musician's SMT was remeasured, and results showed that each musician's SMT did not change compared to the SMT measured before the duet performance [4]. This study analyzing musical performance in a duet setting revealed that the SMT affects the MA in social tasks requiring periodic synchronization (i.e., rowing or playing music), but the task did not affect each musician's SMT [4]. The evidence from all these studies in both solo and duet settings highlights the effect that the SMT has on both the MA and the musical tempo.

Theoretical models have previously proposed mechanisms for the SMT and the MA, but independently. Potential mechanisms of the SMT include motor resonance governed by body weight and limb length [17] and central pattern generators in the brain that are essential for motor control [18, 19]. Mechanisms that have been proposed to explain the MA include delayed recurrent feedback in central-peripheral communication between the auditory and motor systems [20–22] and under- or over-estimation of IOI lengths [23]. Currently, a complete explanation of how the SMT and the MA relate to each other is missing, but the synchronization dynamics of non-linear oscillators can capture some of these relationships. Non-linear oscillator models capture the SMT using an oscillator's natural frequency [3, 24, 25]. Additionally, non-linear oscillators can synchronize in-phase with a periodic stimulus with a frequency that is different but close to the natural frequency. In such a scenario, a MA is observed between the oscillator and the stimulus. Similar to observations made in humans [15], the MA shrinks as the difference between the stimulus frequency and the oscillator natural frequency is reduced [26, 27]. As a result, a non-linear oscillator can capture how a musician's MA grows as a function of the difference between the stimulus tempo and the musician's SMT. Moreover, after a non-linear oscillator is phase-locked to the stimulus frequency, if the stimulus disappears the oscillator will continue oscillating but will return to its natural frequency [26, 27]. This behavior is similar to some extent to musicians' tendency to return to their SMT in the absence of a pacing stimulus [2] or after a musical performance [4]. However, these observations in nonlinear oscillators require the natural frequency and the stimulus frequency to be relatively close to each other, and if the difference exceeds a certain threshold, in-phase synchronization will not be possible and unsynchronized behavior may be observed [26, 27]. This is distinct from the synchronization abilities of humans, who can synchronize with stimuli tempi that are relatively far away from their SMT. Hence, a model consisting purely of non-linear oscillators has a limited potential to fully capture the relationship between the SMT and the MA in humans.

Righetti and colleagues described Hebbian dynamics for frequency learning, which allow an oscillator to adaptively change its frequency to match an external periodic stimulus frequency [28]. A limitation of the model by Rigetti and colleagues is that, unlike humans, the oscillators with frequency learning capabilities do not return to their original natural frequency after stimulation ceases. In contrast, an individual's SMT is recovered after synchronizing with a periodic stimulus with an arbitrary tempo [15]. This phenomenon could be explained by the SMT having an elastic force. This hypothesis proposes that the SMT is an attractor state pulling the system's tempo towards optimal energy usage [3, 15, 29]. Lambert and colleagues added a linear elasticity force governed by Hooke's law to the frequency learning rule described by Righetti and colleagues [30]. This elasticity term pulls the oscillator's frequency to its natural frequency. Hence, the oscillator can match its frequency with an external stimulus frequency using Hebbian learning, but a force is constantly pulling the oscillator's frequency to its natural frequency. Additionally, if the stimulus ceases, the

oscillators frequency will return to the natural frequency.

Behavioral studies have proposed an elastic tempo learning hypothesis to explain the relationship between the SMT and the MA in humans [15]. However, no study so far has quantified how well a dynamical systems model with elastic frequency learning could capture dynamics observed in human data. In this study we test this hypothesis using an oscillatory dynamical systems model with Hebbian learning for frequency adaptation and an elastic force constantly pulling to the natural frequency. More specifically, we use the canonical neural oscillator model described by Large and colleagues [31], and add elastic frequency learning dynamics [28, 30]. The goal is to validate the model using real human data from previous studies of musical performance that analyzed the relationships between the SMT and the MA. If we can achieve either goal, we will show how an individual's SMT affects the MA during coordination with a metronome or another individual. Our modeling approach is appropriate because oscillators have a fixed natural frequency observed spontaneously in the absence of an external stimulus. This natural frequency models the SMT. Because our model has an elastic frequency learning rule, it can adapt its frequency to synchronize with a periodic stimulus of any frequency, but displaying the MA due to the elastic force constantly pulling the oscillator's frequency to its natural frequency. This elastic frequency learning rule simulates two human features. First, the ability to synchronize with a broad range of periodic stimulus tempi. And second, the consistent SMT displayed by humans, even after synchronization with a stimulus of a different tempo.

Given all of its features, we refer to our model as the Adaptive Synchronization with Hebbian Learning and Elasticity (ASHLE). Table 2 presents an overview of the most important ASHLE parameters and their functions. The ASHLE model is inspired by neuroscientific hypotheses about the mechanisms of coordination (references). The methods section gives the complete definition of ASHLE, but here we describe its main properties. ASHLE consists of two oscillators. Both oscillators share the parameters α and β, which lead to spontaneous oscillatory activity. The first oscillator simulates auditory neural entrainment with the stimulus frequency [32–36]. It incorporates a Hebbian rule that learns the frequency of the external stimulus. The second oscillator is coupled to the first oscillator, and simulates entrainment of motor activity to control the action (i.e., fingers playing a piano). This second oscillator includes an elastic Hebbian rule to learn the stimulus frequency (from the first oscillator), while constantly being pulled to its natural frequency. In ASHLE, the timescales of the Hebbian frequency learning rule and the elasticity rule are relatively fast (see the parameter analysis in the methods section). The parameters λ_1 and λ_2 are the stimulus frequency learning rate and the elastic force pulling to the natural frequency, respectively, and act as opposing forces. Additionally, the first oscillator is weakly connected (with strength γ) to the elastic Hebbian rule of the second oscillator, forcing the first oscillator to return to the natural frequency, but at a very slow timescale and only in the absence of a stimulus. γ is very small, and has a negligible effect when a stimulus drives ASHLE. However, in the absence of a stimulus, γ makes ASHLE slowly return to its natural frequency.

ASHLE only simulates the musical beat in a musical performance and not any other spectral features like pitch, harmony or melody content. However, because our goal is to simulate the dynamics of the MA and the SMT at the level of the musical beat, ASHLE is an adequate minimal model. Below we describe three experiments (see Fig 1) we carried out with ASHLE to simulate musician data from three previously published behavioral studies of musical performance.

In our first experiment, we optimized ASHLE to simulate solo musician performance of a simple melody paced by a metronome. This task and data were published by Scheurich and colleagues [15]. In this task, solo musicians performed a melody with a metronome IOI that was 15% or 30% shorter or longer than their SMP (Fig 1A). Their

Table 2. ASHLE parameters and function

Parameter	Value	Function
α	1	bifuraction parameter
β	-1	nonlinear damping
λ_1	4	Hebbian frequency learning rate
λ_2	2	Elastic pull to the natural frequency
γ	0.02	Weak pull to the natural frequency
f		frequency learned Hebbianly
f_0		natural frequency

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). In the musician experiment, the metronome tempo was different from the musician's SMT, and we simulate the same experimental conditions. 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). In the musician experiment, musicians started at a tempo that was different from their SMT. This specific example shows a performance that started with a tempo that was faster than the SMT, and the tempo 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 natural frequency (bottom). (C) The task stimulated in experiment 3, in which pairs of musicians played a simple melody together after hearing four pacing metronome clicks (top). In the musician experiment, pairs of musicians had matching or mismatching SMTs. 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).

data showed that the MA grew as a function of the difference between a musician's SMP and the metronome IOI. We simulated this task using ASHLE. We hypothesized that ASHLE's elastic frequency learning will allow for synchronization with stimuli at any frequency. Additionally, we hypothesized that the elastic force pulling towards ASHLE's natural frequency will cause the MA between ASHLE and the stimulus to grow as the difference between the stimulus frequency and ASHLE's natural frequency.

In our second experiment, we tested whether the same model, with the same parameters, can simulate solo musician performance of a simple melody without a metronome (unpaced). This task and data were published by Zamm and colleagues [2]. In this task solo musicians played a melody, without a metronome, starting at five different spontaneous tempi (Fig 1B). First, at each musician's SMT, then at spontaneous tempi faster and slower than the SMT, and finally at spontaneous tempi that were even faster and even slower than the SMT. Their data showed that musicians have a tendency to return to their SMT after starting a performance at a tempo faster or slower than their SMT. We hypothesize that our model can simulate the same progressive return to the SMT because of the elastic force constantly pulling ASHLE to its natural frequency.

In our third experiment, we tested whether the same model and parameters can simulate duet performance of a simple melody. This task and data were published in another study by Zamm and colleagues [4]. In this task, pairs of musicians played a melody in synchrony after listening to a metronome that established the tempo and

stopped. Musician duets were assigned to either of two experimental groups: matching or mismatching SMTs. The data by Zamm and colleagues [4] showed that the mean absolute asynchrony was larger between musicians with mismatching SMTs than musicians with matching SMTs. We hypothesize that ASHLE's elastic frequency learning will allow two of our models to synchronize with each other independent of whether their natural frequencies are close to each other, but that the asynchrony between two different ASHLE models will grow as a function of the difference between their natural frequencies.

These three musical tasks and results capture relationships between the SMT and the MA. Our goal is to test whether ASHLE's elastic frequency learning can capture these relationships by simulating real musician data. If we are successful, we will shed light on potential mechanisms that give rise to the SMT and the MA. Our model is the first attempt at a dynamical model capturing the relationship between the SMT and the MA that is validated by human data, while also making musician data predictions that can be tested empirically in future experiments. Moreover, due to its dynamical systems nature, ASHLE hypothesizes that the SMT and the MA are related to each other through homeostatic mechanisms at play during PAC.

Results

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

We used ASHLE to simulate the solo task by Scheurich and colleagues [15] 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 data with the mean adjusted asynchrony between the musician beats and the metronome beats across the four different experimental conditions. Scheurich and colleagues [15] reported the mean adjusted asynchrony in their results. The mean adjusted asynchrony is equal to the mean asynchrony during performance with a metronome period different from the SMP, minus the mean synchrony during performance with a metronome period matching the SMP. The musician data shows that the mean adjusted asynchrony was positive (negative) when the metronome period was shorter (longer) than the musician's SMP, and that the asynchrony grows as a function of the difference between SMP and metronome period (see Fig 2A).

We can use ASHLE to simulate the beat during this music performance task. We hypothesized that ASHLE's frequency learning features will allow it to synchronize with a stimulus period shorter or longer than the period associated with its natural frequency, but an asynchrony between ASHLE and the stimulus will be observed due to the elastic force constantly pulling ASHLE's to its natural frequency. This first experiment uses a sinusoidal stimulus to simulate the metronome.

We simulated this task using 20 different ASHLE models that differed in their natural frequencies, which were computed to match each of the 20 musician SMTs measured by Scheurich and colleagues [15] (250ms, 260ms, 300ms, 310ms, 325ms, 340ms, 345ms, 350ms, 380ms, 400ms, 410ms, 430ms, 440ms, 450ms, 460ms, 465ms, 475ms, 480ms, 600ms, and 650ms). Across all 20 ASHLE models simulated in this first experiment, the stimulus was a complex sinusoid $x = \exp(i2\pi f_s t)$, where f_s is the stimulus frequency in hertz, and was different across all simulations. At the beginning of each simulation, ASHLE's frequency variable was set to its natural frequency. The sinusoidal stimulus force was $F = 1$. The frequency learning rate parameter, λ_1, was greater in magnitude than the elastic pull, λ_2, but they were relatively close, resulting

Fig 2. Simulation of the MA between a musician's beat and a metronome beat with a period shorter or longer than the musician's SMP during solo musical performance. (A) The mean adjusted asynchrony (and standard error; N=20) between the musician beat and metronome beat during performance of a simple melody in four conditions: metronome period 30% shorter (F30), 15% shorter (F15), 15% longer (S15), and 30% longer (S30) compared to the musician SMP. The F30, F15, S15, and S30 names for the x axis were used by Scheurich and colleagues [15] to describe the "faster" and "slower" tempo compared to the SMT. Their use is incorrect, however, as a ratio (or percent) of the SMP determines the experimental metronome period, and not a ratio of the speed or rate (implied by the use of 'F' and 'S' for "faster" and "slower", respectively). Here we have pointed out their mistake to clarify the error, but we use their same original labels for the x axis in order to compare our simulation results with their original data. (B) Our simulation results showing the mean adjusted asynchrony (and standard error; N=20) between ASHLE and a sinusoidal stimulus in six conditions: stimulus period 45% shorter (F45), 30% shorter (F30), 15% shorter (F15), 15% longer (S15), 30% longer (S30), and 45% longer (S45) than the period of ASHLE's natural frequency. The shaded bars represent predicted measurements for data that has not been collected yet from musicians. (C) Mean adjusted asynchrony predictions when different ASHLE models (with different natural frequencies) synchronize with a stimulus period that is 45% shorter (F45), 30% shorter (F30), 15% shorter (F15), 15% longer (S15), 30% longer (S30), or 45% longer (S45) than the period of their natural frequecy. The values in (C) are predictions for musician data that has not been collected yet.

in learning of the stimulus tempo (due to λ_1) but with an asynchrony between ASHLE and the stimulus (due to λ_2). In this experiment the parameter γ has a negligible effect because of its small value compared to λ_1, which dominates ASHLE's frequency learning from an external stimulus (see parameter analysis in the methods section for an examination of how different ASHLE parameters values affect its synchronization behavior). Our results show that we have identified the optimal set of parameters that allow ASHLE to capture the mean adjusted asynchrony values observed in the behavioral data by Scheurich and colleagues [15].

For each ASHLE model we simulated the task in four experimental conditions where the stimulus period was 15% or 30% shorter or longer than the period of the natural frequency. We measured the asynchrony between ASHLE and the stimulus (see the methods section for details about our simulation setup, procedure, and measurements). We optimized ASHLE's parameters to approximate the musician data shown in Fig 2A (see model optimization in the methods section). Fig 2B shows the mean adjusted asynchrony that we observed in our simulations for the experimental conditions tested by Scheurich and colleagues [15]. Our simulations show similar results to the human data observed in Fig 2A. Additionally, the shaded bars in Fig 2B show ASHLE's prediction for what the mean adjusted asynchrony would look like if the same group of musicians was to perform with metronome periods 45% shorter or longer than their individual SMP. Our model predicts that the mean adjusted asynchrony in this condition will be even larger compared to when performing with a metronome period 15% or 30% shorter or longer than the musician SMP. Fig 2C shows different ASHLE models (with different natural frequencies) carrying out the same melody performance task with a metronome period that is 45%, 30%, and 15% shorter or longer than their natural period. Individual ASHLE models in Fig 2C predict that the mean adjusted asynchrony grows as a function of the difference between their natural and the stimulus period. Additionally, a longer natural period predicts larger asynchronies across conditions compared to a shorter one. While the shaded bars in Fig 2B are predictions for the group of musicians originally tested, the results in Fig 2C are predictions that

ASHLE makes for individual musician data that has not been collected yet.

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

Experiment 1 showed that our model can capture the MA that Scheurich and colleagues [15] observed when musicians individually performed a simple melody paced by a metronome faster or slower than their SMT. In this second experiment, we used ASHLE to simulate the data by Zamm and colleagues [2], who studied what happens when musicians perform a simple melody unpaced, but starting at a tempo different than the SMT (Fig 1B). Their experiment had four experimental conditions: starting performance tempo fast, faster, slow, and slower than the SMT. Fig 3A shows their results, which consisted of the adjusted slope (mean across musicians) in each of the four different experimental conditions. The adjusted slope is equal to the slope that best fit consecutive IBIs during a melody performance starting at a tempo different than the SMT, minus the slope that best fit the consecutive IBIs during a melody performance that starts at the SMT. The 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; see Fig 3A). Their results show that musicians had a tendency to return to their SMT during a simple melody performance that started at a tempo different than their SMT [2].

Fig 3. Simulation of the slope between consecutive IBIs when an unpaced musician performs a melody starting at a tempo that is different than the SMT. (A) The mean adjusted slope of consecutive IBIs (and standard error; N=24) when solo musicians perform a simple melody starting a tempo that is fast, faster, slow, or slower compared to the musicians SMTs. (B) Our simulation results showing the mean adjusted slope of consecutive IBIs (and standard error; N=23) when ASHLE oscillates, without a stimulus, starting at a frequency that is fast, faster, slow, or slower compared to the natural frequency. (C) Adjusted slope predictions when different ASHLE models (with different periods of natural frequency) oscillate without stimulation, starting with a period that is 45% shorter (F45), 30% shorter (F30), 15% shorter (F15), 15% longer (S15), 30% longer (S30), or 45% longer (S45) compared to the period of ASHLE's natural frequency. For consistency with experiment 1 results, here we also use the F30, F15, S15, and S30 names for the x axis that Scheurich and colleagues [15] use to describe the "faster" and "slower" tempo conditions. The values in C are predictions for musician data that has not been collected yet.

In this second experiment we used the same set of parameter values as in experiment 1. The only values that differed for this experiment were $F = 0$ (due to the lack of stimulus), and the natural frequency and initial frequency values, which were dictated by human data. In this experiment, the Hebbian frequency learning parameter λ_1 has no effect because there is no stimulus. The elastic pull to the natural frequency, λ_2, is in action, as well as the weak pull to the nautral frequency in perception, $\gamma = 0.02$. λ_2 and γ act as additive forces that control the size of consecutive period lengths in ASHLE's behavior when a stimulus is not present. However, while λ_2 dictates how strongly ASHLE pulls its frequency toward its natural frequency, γ dictates how quickly ASHLE forgets its frequency in favor of the natural frequency. Here, γ is a small value, keeping ASHLE close to its frequency and only letting ASHLE slowly return to its natural frequency. That explains why the resulting slope values in Fig 3B and Fig 3C are

relatively small (see parameter analysis in the methods section for an examination of how different ASHLE parameters values affect its behavior).

We used ASHLE to simulate the beats during this music performance task. We hypothesized that ASHLE's pulling force will make it return to the natural frequency, but at a very slow rate due to the weak pull to the natural frequency (small γ parameter; see the methods section for a complete mathematical description of ASHLE and parameter analysis). In contrast with experiment 1, in this second experiment ASHLE was never stimulated by a sinusoid. We simulated this task using 23 ASHLE models that differed in their natural frequency, computed from the 23 musician SMP values measured by Zamm and colleagues [2] (320ms, 350ms, 355ms, 359ms, 382ms, 390ms, 390ms, 415ms, 418ms, 430ms, 435ms, 438ms, 439ms, 439ms, 443ms, 445ms, 455ms, 457ms, 462ms, 470ms, 475ms, 525ms, and 572ms; we excluded the data from the participant with the SMP of 665ms due to the large difference between this participant's spontaneous rates and the rest of the group; see the methods section for a detailed description of our rationale to exclude this participant). For each ASHLE model we simulated the task in four experimental conditions where ASHLE's initial frequency was fast, faster, slow, or slower compared to the natural frequency (see the methods section for details about our simulation setup and procedure). These four initial frequency values matched the measurements that Zamm and colleagues [2] made for each musician performing at a tempo that was spontaneously fast, faster, slow, or slower than the SMT. Fig 3B shows the adjusted slope that we observed in our simulations for the four different experimental conditions tested by Zamm and colleagues [2]. Our simulations show similar results to the human data observed in Fig 3A. Additionally, Fig 3C shows musician data predictions by simulating different ASHLE models (with different nautral periods) that perform the same melody task with an initial period that is 45%, 30%, and 15% shorter or longer than the natural period. Individual ASHLE models in Fig 3C predict that the slope grows as a function of the difference between initial period and the natural period. Additionally, a slower natural frequency predicts larger slope values compared to a faster one. The results in Fig 3C are predictions that ASHLE makes for musician data that has not been collected yet.

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

Experiments 1 and 2 showed that the same ASHLE model parameters can simulate two different behavioral measurements collected in two independent experiments: the MA in paced musical performance and the slope of consecutive IBIs in unpaced musical performance. In this third experiment, we used ASHLE to simulate another task by Zamm and colleagues [4] 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.

We used the same ASHLE parameters as in experiments 1 and 2., In addition, $F_z = 0.01$, the coupling strength between ASHLE models stimulating each other, and the addition of gaussian noise to ASHLE's action oscillator to simulate noise, leading to the absolute asynchrony magnitudes observed in the behavioral data (see model optimization in the methods section). Without the gaussian noise, ASHLE showed the same relative difference in the mean absolute asynchrony between experimental groups (match or mismatch SMT), but the values were considerably smaller (see model optimization in the methods section). We simulated this task using 20 pairs of ASHLE models that differed in their natural frequencies, computed from the 40 musician SMTs measured by Zamm and colleagues [4]. We paired ASHLE models replicating Zamm and colleagues' [4] pairing of musicians with matching or mismatching SMTs. For each

pair we simulated the duet melody performance task, where musician duets performed the same simple melody four consecutive times (see the methods section for details about our simulation setup and procedure). Fig 4B shows the mean absolute asynchrony that we observed in our simulations for the experimental groups and melody repetitions tested by Zamm and colleagues [4]. Our simulations show similar results to the human data observed in Fig 4A. Additionally, Fig 4C shows musician data predictions by simulating different ASHLE models performing the same duet melody task with another ASHLE model that has an period of natural frequency 220ms shorter, 110ms shorter, 10ms shorter, 10ms longer, 110ms longer, and 220ms longer than the other ASHLE model in the duet. The results in Fig 4C predict that the mean absolute asynchrony between musician pairs grows as a function of the difference between their natural frequency values.

Fig 4A shows their results, which consisted of the mean absolute asynchrony between the beats of the two performing musicians in each of the two experimental groups, separately for each of the four melody repetitions (see the methods section for a detailed description of the task by Zamm and colleagues [4]). This data shows that the mean absolute asynchrony was smaller between musician duets with matching SMTs compared musician duets with mismatching SMTs. We used ASHLE to simulate the beats during this musical performance task. We hypothesized that two ASHLE models will be able to synchronize with each other due to their frequency learning mechanism. However, the elasticity pulling to their respective natural frequency will result in asynchrony between them, and the asynchrony will be smaller between two synchronizing ASHLE models with similar natural frequency values compared to two ASHLE models with dissimilar natural frequency values. In this third experiment pairs of ASHLE models are stimulated first by a sinusoid that establishes a common frequency and then the two ASHLE models stimulate each other. As a result, stimulation in this third experiment is slightly more complex than the constant sinusoidal stimulation in experiment 1 and the lack of stimulation in experiment 2.

Fig 4. Simulation of the mean absolute asynchrony between two musicians with matching or mismatching SMTs during duet musical performance. (A) The mean absolute asynchrony (and standard error; in each experimental group N=10) between two musicians with matching or mismatching SMTs during performance of a simple melody four consecutive times. (B) Our simulation results showing the mean absolute asynchrony between two synchronizing ASHLE models with natural frequency values that are close or far from each other. (C) Mean absolute asynchrony predictions when different ASHLE models (with different periods of natural frequency) synchronize with another ASHLE model with a period of natural frequency that is 220ms shorter, 110ms shorter, 10ms shorter, 10ms longer, 110ms longer, and 220ms longer. The values in C are predictions for data that has not been collected yet from musicians.

Discussion

The interaction of Hebbian learning and elasticity mechanisms during beat tracking can capture various coupling dynamics

In experiments 1 and 3, the asynchrony between ASHLE and a periodic stimulus was affected by the elastic Hebbian frequency learning rule. Similarly, experiment 2 showed that, in the absence of a stimulus, the elastic Hebbian frequency learning also affected the frequency of ASHLE's spontaneous activity, especially if ASHLE's frequency was different than its natural frequency. The rate of ASHLE's oscillatory behavior is the

result of elastic frequency learning, with Hebbian frequency learning and the elasticity acting as opposing forces. While Hebbian frequency learning lets ASHLE change its frequency to match any stimulus frequency, the elastic force pulls ASHLE toward its natural frequency. ASHLE's behavior is consistent with theoretical accounts highlighting that a dynamical system's natural frequency is the optimal state for synchronization, even if the system can synchronize at other frequencies [15, 37–39]. An interesting question is where in the musician brain these mechanisms of frequency learning and frequency elasticity occur. As far as frequency learning, previous research shows that cortical and subcortical networks in motor brain areas synchronize neural activity that reflects the rhythms of a periodic stimulus [32, 35, 40, 41]. The origins of the SMT in humans (simuated by ASHLE's natural frequency) involve both the central and the peripheral nervous system [18] including anatomical and biomechanical factors [17]. Currently no consensus exists about the origins of the SMT and its influence during PAC. ASHLE is the first model validated by real human data that captures how motor actions entrain to the tempo of a periodic stimulus and are influenced by the dynamic principles of elastic Hebbian frequency learning. By using a dynamical systems model, we rule in a family of possible biomechanical principles [42] that could underlie the origins of the SMT.

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

Our model was able to reproduce the mean adjusted asynchrony between musician beats and metronome beats when the metronome tempo is faster or slower than the musician's SMT, in a simple musical performance task (Fig 2A and Fig 2B). Taking a closer look at the behavioral data in Fig 2A we see that the magnitude of the mean adjusted asynchrony values is slightly asymmetric between faster and slower stimuli with respect to the SMT. Faster metronomes resulted in slightly larger mean adjusted asynchrony magnitudes than slower ones. This human behavior implies frequency-scaling of the elastic Hebbian learning rule in Eqs (5b) and (5d) (see [26, 31] for more theory about frequency-scaled dynamical systems similar to ASHLE). Additionally, ASHLE's elasticity term $-\lambda_2 \left(\exp\left(\frac{f_m - f_0}{f_0}\right) - 1\right)$ in Eq (5d) is an exponential function so that a stimulus faster than ASHLE's natural frequency (causing $\exp\left(\frac{f_m - f_0}{f_0}\right) - 1 > 0$) would result in an exponentially larger force toward the natural frequency compared to a stimulus slower than ASHLE's natural frequency (causing $\exp\left(\frac{f_m - f_0}{f_0}\right) - 1 < 0$). Due to frequency scaling, higher values of f_e and f_m amplify the learning rate (λ_1) and the elastic force (λ_2), while lower values of f_e and f_m shrink them. In ASHLE f_e and f_m are always very close (if not identical) in value to each other. ASHLE's mechanisms and behavior are validated by its ability to simulate the data by Scheurich and colleagues [15] (Fig 2A and Fig 2B).

ASHLE's mechanisms of frequency learning and elastic force pulling to the natural frequency are also conceived in neuroscientific terms. ASHLE's frequency learning rate parameter, λ_1, simulates the brain's ability to entrain perceptual networks when processing a periodic stimulus like a musical beat or a metronome tempo [32]. In contrast, ASHLE's elasticity parameter, λ_2, reflects the neural and biomechanical properties that give rise to the SMT [17]

While ASHLE was able to simulate the musician data by Scheurich and colleagues [15], the model used in experiment 1 does not show variability in its dynamics. Once ASHLE phase-locks with the sinusoidal stimulus, the asynchrony between ASHLE and the stimulus reaches a steady state with zero variability. This is a limitation of ASHLE's ability to simulate musician's beat tracking behavior, reflected by

a degree of variability across musician IBIs. In experiment 1, this lack of variability was a limitation to perfectly simulate the results by Scheurich and colleagues [15]. Using ASHLE models with different natural frequency values allowed us to simulate the variability between musicians (hence the standard error in Fig 2B). However, our standard errors were overall smaller compared to the musician standard errors that Scheurich and colleagues [15] reported. The standard error in musicians has two sources: (1) variability within each musician's beat tracking behavior and (2) variability between the behavior of different musicians. While we were able to simulate the variability between musicians using ASHLE models with different natural frequency values, the lack of noise in ASHLE's activity is the likely cause of the difference between our simulation results and the musician results. Adding gaussian noise to ASHLE's activity could better approximate the musician results. However, we considered the addition of noise to be beyond the scope of this first experiment because our goal was to optimize ASHLE to simulate the magnitude of the mean adjusted asynchrony (see model optimization in the methods section).

ASHLE also made testable predictions of paced musician performance in a solo setting (Fig 2B and Fig 2C). First, it predicted mean adjusted asynchrony values that we could expect 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 2B shaded bars). Additionally, Fig 2C made 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 the SMT. These simulations sho that ASHLE can be used to make predictions about a musician's behavior with any SMT, and such predictions can be empirically tested to further validate ASHLE's behavior or better tune its parameters.

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

Our model was also able to simulate the mean adjusted slope between consecutive IBIs that Zamm and colleagues [2] observed when musicians performed a melody without a metronome, starting at a tempo faster or slower than the SMT (Fig 3A and Fig 3B). ASHLE's approximation of the data is not perfect, but it is close. This is because the parameters λ_1 and λ_2 were optimized to simulate the data in experiment 1 and the same values were used in experiment 2. Hence, our results in this second experiment show that the values for λ_1 and λ_2 can generalize to simulate data across different datasets. There was only one ASHLE parameter that we optimized based on the musician data from experiment 2. This was the γ parameter (see model optimization of experiment 2, in the methods section). In experiment 2 there is no sinusoidal stimulus ($F = 0$) and γ in Eq (5a) controls how quickly ASHLE's frequency returns to the natural frequency. $\gamma = 0.02$ was the value that resulted in the best fit between ASHLE's simulation results and the musician data results. The small $\gamma = 0.02$ value can be explained by looking at the musician data in Fig 3A. Musician's mean adjusted slope values are relatively small for all experimental conditions. These small values result from musicians showing a tendency to very slowly return to their SMT throughout the musical performance. That is why the value of γ is very small compared to $\lambda_1 = 4$ and $\lambda_2 = 2$. Hence, γ is the ASHLE parameter that represents a musician's tendency to return to their SMT in the absence of stimulation. It is relatively small because part of musical training consists in being able to maintain an arbitrary tempo throughout a musical performance [43, 44].

Theoretical models of musical tempo suggest entrainment of periodic neural activity with periodic actions [45], but the mean adjusted slope results by Zamm and colleagues [2] also suggests that the motor system, while maintaining the musical tempo,

is constantly influenced by biomechanical and neural constraints that could give rise to the SMT [17]. Evidence from synchronization-continuation tasks suggest the existence of a physical and likely anatomical force pulling human periodic actions to the SMT in the absence of an external stimulus [3,16]. Our simulations reflect that the same mechanisms to maintain a tempo are at play in experiment 1 and 2. However, the presence of a stimulus in experiment 1 resulted in maintaining the tempo and the observation of an MA between the musician and the metronome. In contrast, the lack of stimulus in experiment 2 results in the musician slowly returning to the SMT.

As in experiment 1, the ASHLE model from experiment 2 was deterministic. The standard errors in our simulation results (Fig 3B) are different from the standard errors in the musician results (Fig 3A). As we discussed in experiment 1, the results have two sources of activity: variability in each individual musician's actions and variability between musicians. Our simulations only capture the variability between musicians by using different natural frequency values to simulate different musician's SMTs. ASHLE's lack of variability in its activity suggest that the variability in each individual musician's actions also contributes to the standard error observed in Fig 3. Our simulations of experiment 2 show that ASHLE is a good model to approximate the slope between consecurive IBIs, but not the variability observed when musicians try to maintain the tempo in a solo setting without a pacing metronome.

ASHLE also made testable predictions of unpaced musician performance in a solo setting. Fig 3C shows predictions at the individual musician level simulating the mean adjusted slope that musicians with different SMT would produce when performing a melody starting at various tempi that are faster or slower than the SMT. ASHLE can be used to make predictions about a musician with any SMT, and such predictions can be empirically tested to further validate ASHLE's behavior.

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

ASHLE was able to simulate the mean absolute asynchrony between the beat of musician duets (with matching or mismatching SMTs) that performed a simple melody (Fig 4A and Fig 4B). With the exception of the type of stimulation (which was different between all three experiments) we used the same ASHLE model parameters in experiments 1, 2, and 3. By carrying out experiment 3, our goal was to show that ASHLE can capture musical performance behavior not just in solo settings, but also in a duet setting. To capture the data of experiment 3, it was necessary to add noise to ASHLE's action oscillator (see model optimization in the methods section). Without the gaussian noise, ASHLE showed the same qualitative pattern of results, but with a noticeable difference in magnitude between the musician mean absolute asynchrony and ASHLE's (see Fig 5). Adding noise to ASHLE's effector significantly improved ASHLE's approximation of the human data (Fig 4A and Fig 4B). More specifically, the added noise results in ASHLE a better approximation of the magnitude of the mean absolute asynchrony in the musician duets. Across all simulations, the standard error in ASHLE's results is very similar to the standard error displayed by the human data. While we could have changed parameter values to achieve a perfect match between ASHLE's simulations and the musician activity for all experimental conditions, our goal was to identify the best set of parameter values that could capture the activity across different musician experiments.

ASHLE was able to simulate the solo performance results in experiments 1 and 2 with a deterministic model. However, noise was necessary to simulate the duet performance results in experiment 3. This makes sense as the musical performance task in experiment 3 involves more sources of variability than the tasks in experiments 1 and

Fig 5. Simulation without noise of the mean absolute asynchrony between two musicians with matching or mismatching SMTs during duet musical performance. Our simulation results showing the mean absolute asynchrony between two synchronizing ASHLE models with matching or mismatching natural frequency, but no noise added to Eq (5c). The resulting mean absolute asynchronies in this simulation without noise are much smaller compared to the musician data results in Fig 4A. The added noise in Eq (6) improves the model's results, which are shown in Fig 4B.

2. While only one musician performs in experiments 1 and 2, in experiment 3 there are two performing musicians receiving feedback from each other. The variability in musician behavior noticeably influences the absolute asynchrony between two musicians performing as a duet. ASHLE's need for noise in its action oscillator reveals that the variability in musician behavior at the individual level is a significant contributor to the mean absolute asynchrony in experiment 3. Without noise, ASHLE simulated a mean absolute asynchrony around 2ms and 18ms for the duets with matching and mismatching SMTs, respectively. When noise was added, ASHLE showed a mean absolute asynchrony around 10ms and 20ms for the duets with matching and mismatching SMTs, respectively. Hence, the presence of noise is needed to capture musical duet behavior, especially when the two musicians have matching SMTs.

To simulate this musical duet task, we also needed a significantly smaller connection strength (see model optimization in the methods section) between the two ASHLE models stimulating each other. This difference in stimulus strength between experiment 1 and experiment 3 is due to the nature of the stimulus. With unidirectional forcing, a large forcing strength helps achieve phase-locked synchronization between a dynamical system and a stimulus [26]. This is very different when two or more dynamical systems stimulate each other, like the two ASHLE models stimulating each other in experiment 3. Previous studies of stability between dynamical systems have revealed that small forcing strengths between dynamical systems ensure the stability of their periodic behavior [26].

ASHLE also made testable predictions about how musician duets would perform in new experimental manipulations. Fig 4C shows predictions for the mean absolute asynchrony that one would observe between two musicians, one with a specific SMT, and the other with an SMT that differs by a specific value from the other musician's SMT. These predictions show 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. ASHLE can be used to make predictions about musician duets with any SMTs, and such predictions can be empirically tested to further validate ASHLE's behavior.

General discussion

We have presented the ASHLE model, which captures how the musical tempo is maintained during various simple performance tasks. In experiment 1, ASHLE captured the MA between a musician's beats and metronome beats during a simple melody performance task (Fig 2). In experiment 2, ASHLE captured the rate of beat change when a solo musician performed a simple melody without a metronome (Fig 3). And in experiment 3, ASHLE captured the mean absolute asynchrony between the beat of two musicians performing a simple melody as a duet (Fig 4). ASHLE was able to capture these three different tasks using the same set of parameters for its intrinsic dynamics (see parameter analysis and model optimization in the methods section). Our simulations validate ASHLE's ability to simulate real musician data. Hence, ASHLE is a model with a mechanisms of elastic Hebbian frequency learning that generalizes across different tasks and different datasets. We also used ASHLE to make testable predictions

of musician performance data not yet collected (Fig 2B, Fig 2C, Fig 3C, and Fig 4C). Future behavioral studies can collect this data to validate ASHLE as a model able to capture musician performance. More data collected could also help to better fine tune ASHLE's parameters in case ASHLE, in its current state, cannot sufficiently predict new musician data.

ASHLE is the first model with the interacting forces of Hebbian frequency learning and elasticity pulling toward a systems natural frequency to be applied to asynchrony data. Previous models have focused on explaining the origin of the negative mean asynchrony (NMA), which is a specific type of asynchrony when the synchronizing agent's actions anticipate the pacing signal (i.e., a metronome). The first kind of these models propose that the NMA is the result of adaptive synchronization by means of phase and period correction rules at phenomenological [46] and neuromechanistic levels [42]. Another kind of model, called strong anticipation models, propose that the NMA is the result of delayed feedback [20] and, in the case of humans, the NMA has been theorized to emerge from delayed neural communication in the sensorimotor system [21]. ASHLE can capture both the NMA and the positive MA using the same fundamental principles of adaptive frequency learning and elastic force. In humans, the NMA is observed when synchronizing with a metronome tempo with an IOI larger than 350ms [9]. Interestingly, the average human SMP is around 400ms. As a result, the human NMA in ASHLE emerges from a force pulling the human tempo to the SMT during adaptive synchronization with a slower stimulus. However, there are similarities between ASHLE and models of strong anticipation to explain the NMA as resulting from delayed feedback [20, 21]. Both ASHLE and strong anticipation models are based on physical principles. Future work should try to combine ASHLE and the strong anticipation model of human anticipation to achieve a comprehensive model of adaptive and anticipatory activity in human behavior.

We have presented a versatile, neuroscientifically-inspired, and ecologically valid model of adaptive musical beat performance for solo and duet settings. ASHLE is a model that can capture adaptive human synchronization based on dynamical principles. ASHLE could be used not just to simulate human data and predict human behavior, but also as a tool with which humans can interact in experimental, musical, and therapy settings. ASHLE is the first model of its kind, and provides us with a better understanding about how Hebbian frequency learning in the motor cortex interacts with the elastic and dynamical constraints that give rise to the SMT.

Methods

Model definition

Theoretical background

A non-linear oscillator with a natural frequency f_1 can synchronize with another oscillator or an arbitrary periodic signal of frequency f_2 if $|f_1 - f_2| < \epsilon$, where ϵ is small compared to f_1 and f_2. Synchronization also requires the synchronizing oscillator (or oscillators) to have properly defined parameter values because if any of these parameters is outside the oscillator's phase-locking regime, synchronization will fail [47]. To overcome the limitation of fixed parameters, studies have shown that oscillator parameters can be defined dynamically by a differential equation [45, 47–53]. Because the natural frequency is critical to determine an oscillator's phase-locking regime, Righetti and colleagues [47] defined an equation that allows a limit-cycle oscillator to dynamically adapt its frequency to match the rate of periodic activity of an arbitrary stimulus and phase-lock.

$$\dot{f} = -\lambda x(t)\sin(\phi) \tag{1}$$

In Eq (1), f is an oscillator's natural frequency in hertz, λ is the learning rate (or the adaptation timescale for f), $x(t)$ is an arbitrary stimulus, and ϕ is the oscillator's phase [47]. Because f is a real-valued term, the stimulus $x(t)$ is assumed to be a real signal, but if it were complex-valued, only the real part of $x(t)$ would be needed to observe the adaptive frequency behavior (the imaginary part can also be used. See the ASHLE model definition for this equation). Hence, to process a complex-valued stimulus, Eq (1) would have the following form:

$$\dot{f} = \lambda \Im\{x(t)e^{-i\phi}\} \tag{2}$$

Eq (2) allows a non-linear oscillator to dynamically adapt its frequency to phase-lock with a periodic stimulus (by synchronizing only with its real part) of an arbitrary frequency [47]. However, Eq (2) does not have memory of its initial conditions for the frequency term, so after stimulation the oscillator's frequency f will remain in its learned state and will not return to the oscillator's original natural frequency. Lambert and colleagues [30] envisioned an oscillator able to adapt its frequency elastically so that the oscillator would return to its natural frequency in the absence of stimulation. Eq (3) describes this elastic frequency learning rule:

$$\dot{f} = \lambda_1 \Im\{x(t)e^{-i\phi}\} - \lambda_2 \frac{f - f_0}{f_0} \tag{3}$$

In Eq (3), λ_1 is the frequency learning rate from Eq (2) and λ_2 is the elastic force pulling the oscillator's f back to f_0, which is a fixed value representing the oscillator's natural frequency. λ_2 multiplies the difference between f and f_0. Thus, the new term $-\lambda_2 \frac{f-f_0}{f_0}$ is equivalent to Hooke's law, which states that the force needed to extend a spring is linearly scaled by the distance resulting from such extension [30]. As f and f_0 differ in value due to a stimulus $x(t)$ that changes the adaptive frequency term $-\lambda_1 \Re\{x(t)\sin(\phi)\}$, the system's elastic energy demand increases. Only the stimulus $x(t)$ can supply this energy demand, and in the absence of a stimulus, the system will make $f = f_0$, which is the state with the lowest elastic energy demand. Lambert and colleagues [30] used Eq (3) to create networks of oscillators with different f_0 values. This network allowed them to dynamically find pulse and meter in musical signals [30].

To use Eq (3) we need to combine it with an oscillatory dynamical system. Because we want to simulate periodic synchronization, we use the canonical Hopf oscillator described by Large and colleagues that can synchronize with an external periodic stimulus [26, 31]. The oscillator that we will use has the following general form:

$$\frac{1}{f}\dot{z} = z\left(\alpha + i2\pi + \beta_1|z|^2 + \frac{\epsilon\beta_2|z|^4}{1-\epsilon|z|^2}\right) + x(t) \tag{4}$$

In Eq (4) z is the state of the oscillator, and α, β_1 and β_2 are fixed parameters that control the dynamics of the oscillator. f determines the rate of oscillation and ϵ is a parameter that controls the degree of higher-order nonlinear activity in the oscillator. $x(t)$ is an arbitrary stimulus. This oscillator is derived from a biological model of neural oscillation and can capture neural synchronization observed in real neurons and sensory processing [54, 55]. Moreover, a single oscillator can show stable oscillatory activity, even after it is no longer being stimulated, hence showing "memory" [26]. We want to simulate a musician's SMT, so we only need one oscillator with a fixed f_0 value. This contrasts with the models described by Large and colleagues [31], who used a network of oscillators with different natural frequency values.

The ASHLE model

Eqs (5a), (5b), (5c), and (5d) show the ASHLE model:

$$\frac{1}{f_p}\dot{z}_p = z_p\left(\alpha + i2\pi + \beta|z_p|^2\right) + Fx(t) \tag{5a}$$

$$\dot{f}_p = f_p\left(\lambda_1 \Im\left(Fx(t)e^{-i\phi_p}\right) - \gamma\left(\exp\left(\frac{f_p - f_a}{f_a}\right) - 1\right)\right) \tag{5b}$$

$$\frac{1}{f_a}\dot{z}_a = z_a\left(\alpha + i2\pi + \beta|z_a|^2\right) + \exp(i\angle z_p) \tag{5c}$$

$$\dot{f}_a = f_a\left(\lambda_1 \sin(\phi_p - \phi_a) - \lambda_2\left(\exp\left(\frac{f_a - f_0}{f_0}\right) - 1\right)\right) \tag{5d}$$

Eqs (5a) and (5c) are oscillators like the one in Eq (4), but with $\beta_1 = \beta$, $\beta_2 = 0$, $\epsilon = 0$. Eqs (5a) and (5b) have a subscript p that stands for "perception", while Eqs (5c) and (5d) have a subscript a that stands for "action". In all of our simulations $\alpha = 1$ and $\beta = -1$ so that the intrinsic dynamics of Eqs (5a) and (5c) are those of limit-cycle oscillators. In the absence of stimulus (i.e., when $F = 0$), Eqs (5a) and (5c) will show spontaneous and perpetual oscillation [26]. Eqs (5b) and (5d) are the frequency learning equations for Eq (5a) and Eq (5c), respectively. Eq (5b) has a frequency learning terms $-\lambda_1 \Re\left(iFx(t)\exp(-i\angle z_p)\right)$ and a slow term $-\gamma\left(\exp\left(\frac{f_p - f_a}{f_a}\right) - 1\right)$. The first one learns the frequency of the external stimulus $Fx(t)$, while the second one allows for a slow return to the natural frequency because the parameter γ always has a value smaller than λ_1 and λ_2. So the slow timescale in (5b) is negligible in the presence of a stimulus $Fx(t)$. This slow term is only important when $F = 0$. The two frequency learning terms in Eq (5b) control learning of the stimulus frequency and forgetting of an entrained frequency, respectively (see parameter analysis in the next subsection). Eq (5d) is similar to the one described by Eq (3), which contained a frequency learning and an elasticity term. However, in contrast with Eq (3), Eq (5d) has an exponential elasticity term to better cope with the exponential nature of the hertz scale of terms f_a and f_0. The first term in Eq (5d) carries out entrained frequency tracking while the second one is the pulling force moving ASHLE's frequency toward the natural frequency. Eqs (5b) and (5d) share the λ_1 parameter. While the value of f_0 may change between different simulations that we carried out, the values of λ_1 and λ_2 are always fixed across all our simulation experiments (see parameter analysis and model optimization in the upcoming sections to understand how we picked the values for these two parameters).

This model is inspired by neuroscientific hypotheses about processes involved during perception-action coordination [32]. Eqs (5a) and (5b) simulate entrainment in cortico-subcortical motor brain areas when processing a periodic stimulus of an arbitrary frequency [34–36]. That is why Eq (5a) shows sustained oscillatory activity even after a stimulus is no longer present, simulating neural entrainment to the musical beat [32]. Eqs (5c) and (5d) simulate the motor system processing the entrained beat information to control peripheral effectors (i.e., a finger playing a piano) during a musical performance task. Eqs (5b) and (5d) simulate the frequency learning rules of central and peripheral motor function. Because of how ASHLE processes a stimulus, Eqs (5a) and (5b) simulate how the neurons entrain to a periodic stimulus, while Eqs (5c) and (5d) captures how the motor system converts this encoded information into a motor command to effectors.

While Eq (5a) is driven by an external stimulus $Fx(t)$, Eq (5c) is driven by a unit-magnitude complex sinusoid with the angle of Eq (5a). The input $x(t)$ to ASHLE in Eq (5a) will be a complex sinusoid to simulate musical performance paced by a metronome, zero to simulate spontaneous musical performance, or the state of another

ASHLE's z_a term to simulate duet musical performance using two synchronizing ASHLE models.

Parameter analysis

Eqs (5a), (5b), (5c), and (5d) describe the ASHLE model. Eqs (5a) and (5c) are the Hopf oscillator described by Large and colleagues [31]. Previous studies have investigated how such an oscillator synchronizes with a periodic external stimulus in a phase-locked fashion [26]. For that reason, we focus on analyzing how the parameters λ_1, λ_2 and γ in Eqs (5b) and (5d) affect synchronization with a periodic stimulus. Throughout this chapter, we fixed parameters $\alpha = 1$ and $\beta = -1$. This choice of parameter values causes Eqs (5a) and (5c) to show spontaneous unit-magnitude limit-cycle behavior [26]. In our first and second analysis we use values of $F = 1$ (like in Experiment 1) and in our third analysis $F = 0$ (like in Experiment 2). The specific value of $F_z = 0.01$ that we used in Experiment 3 is described in the model optimization section specific to the Experiment 3 methods section. In the first and second analysis, the external stimulus $x(t)$ was always a unit magnitude complex sinusoid, so $x(t) = \exp(i2\pi f_s t)$, where f_s is the stimulus frequency in hertz and t is time. For all simulations in this analysis, ASHLE's natural frequency $f_0 = 2.5$ because the average musician SMP is around 400ms in previous behavioral studies ($f_0 = \frac{1000\text{ms}}{400\text{ms}} = 2.5\text{Hz}$) [2, 4, 15].

We first analyzed how λ_1 and λ_2 affect the phase-locked asynchrony between the stimulus $x(t)$ and ASHLE's z_a (which simulates control of peripheral effectors) when the stimulus period is 45% shorter ($f_s = \frac{1000\text{ms}}{0.55\times 400\text{ms}}$) and 45% longer ($f_s = \frac{1000\text{ms}}{1.45\times 400\text{ms}}$) than ASHLE's period of natural frequency. We ran different simulations, each with a unique value for λ_1 and λ_2. In all simulations, ASHLE and $x(t)$ interacted for 50 seconds with initial conditions $z_p(0) = z_a(0) = 0.001 + i0$. In Eq (5b), the parameter $\gamma = 0$ because we do not want to study the effect of this small parameter in this first analysis (see our third parameter analysis below to understand how *gamma* affects ASHLE's behavior). After each simulation finished, we found the location (in milliseconds) of local maxima (i.e., peaks) in the real part of the oscillatory activity of $x(t)$ and z_a (Fig 6A). Then, we subtracted the location of the peaks of $x(t)$ from the location of the peaks of z_a and averaged the result to obtain the mean asynchrony in milliseconds. If $x(t)$ and z_a showed a different number of peaks, that indicated that phase-locked synchronization did not occur between the two. Fig 6B and Fig 6C show the result of this analysis as a function of λ_1 and λ_2 when f_s (the frequency of the stimulus) has a period 45% shorter ($f_s = 4.5$; Fig 6B) and 45% longer ($f_s = 1.72$; Fig 6 than ASHLE's period of natural frequency. This analysis revealed that phase-locked synchronization does not occur for certain combinations of λ_1 and λ_2 values (black cells in Fig 6B and Fig 6C). Interestingly, when $\lambda_1 = 0$, phase-locked synchronization is never possible. This makes sense, since λ_1 is the frequency learning rate that allows ASHLE to change its frequency to match the frequency of $x(t)$. When $\lambda_1 = 0$, phase-locked synchronization can only occur between ASHLE and a stimulus with a frequency that is close to ASHLE's natural frequncy (see the study by Kim and Large [26] for a detailed analysis of this kind of synchronization in the absence of a frequency learning rule). This analysis also revealed that as the value of λ_2 becomes larger, phase-locked synchronization may not be observed (depending on the value of λ_1) because ASHLE's frequency terms f_p and f_a are pulled strongly to the natural frequency f_0. The values of λ_1 and λ_2 modulate the size of the asynchrony. As λ_1 becomes larger, the magnitude of the asynchrony between z_a and $x(t)$ tends to decrease and is sometimes close to zero when $\lambda_2 = 0$. The opposite occurs when λ_2 grows. This observation reveals that λ_1 and λ_2 work in opposing directions. While λ_1 changes the model's frequency to match the stimulus frequency, λ_2 pulls ASHLE's to f_0. Moreover, as ASHLE's frequency deviates from f_0, λ_2 acts with

more strength, while the strength of λ_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_a and $x(t)$ is affected by whether $x(t)$ is faster or slower than ASHLE's natural frequency, with a tendency be positive (lagging behavior) and negative (anticipatory behavior), respectively.

Fig 6. The asynchrony between ASHLE and a sinusoid with period 45% shorter or longer than ASHLE's period of natural frequency, as a function of frequency learning and elasticity parameters. (A) Illustration of the asynchrony between ASHLE's z_a and the external sinusoidal stimulus, and how it is measured. (B) The asynchrony in milliseconds between ASHLE and a sinusoidal stimulus with a period 45% shorter than ASHLE's period of natural frequency, and its change as a function of λ_1 and λ_2, which are ASHLE's parameters for frequency learning and elasticity, respectively. (C) The asynchrony in milliseconds between ASHLE and a sinusoidal stimulus with a period 45% shorter than ASHLE's period of natural frequency, and its change as a function of λ_1 and λ_2 values. Black cells indicate λ_1 and λ_2 value pairs for simulations where ASHLE and the sinusoidal stimulus could not synchronize.

Our second analysis focused on narrowing down our search for parameter values for λ_1 and λ_2 that result in asynchronies in the range observed in human data. In the experiment by Scheurich and colleagues [15], musicians synchronized with a stimulus period that is 15% or 30% shorter or longer than their SMP to measure the mean adjusted asynchrony (which is equal to the MA observed during performance with a metronome tempo different from the musician SMT, minus the MA measured during performance with an metronome matching the musician SMT). Musicians showed mean adjusted asynchrony values in the range of -10ms and 10ms (see Fig 2A). In this analysis 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 first analysis) and a sinusoidal stimulus $x(t) = \exp(i2\pi f_s t)$ with six potential period lengths: 45% shorter ($f_s = \frac{1000\text{ms}}{0.55\times 400\text{ms}}$), 30% shorter ($f_s = \frac{1000\text{ms}}{0.70\times 400\text{ms}}$), 15% shorter ($f_s = \frac{1000\text{ms}}{0.85\times 400\text{ms}}$), 15% longer ($f_s = \frac{1000\text{ms}}{1.15\times 400\text{ms}}$), 30% longer ($f_s = \frac{1000\text{ms}}{1.30\times 400\text{ms}}$), and 45% longer ($f_s = \frac{1000\text{ms}}{1.45\times 400\text{ms}}$) than ASHLE's period of natural frequency. In our first parameter analysis, we noticed that values for λ_1 between 3 and 5, and λ_2 between 1 and 3 could result in MA values in the range of the mean adjusted asynchrony observed in humans (Fig 6B and Fig 6C). In this analysis we refine our search for λ_1 and λ_2 in this range of values. We ran simulations in a similar fashion to our first analysis. ASHLE and the sinusoidal stimulus interacted for 50 seconds. ASHLE had initial conditions $z_p(0) = z_a(0) = 0.001 + i0$ and $f_p(0) = f_a(0) = f_0$. In Eq (5b) the parameter $\gamma = 0$ because we want this term to be negligible when a stimulus is present. For each simulation, we found the location (in milliseconds) of the local maxima in the real part of the oscillatory activity of ASHLE's z_a and $x(t)$. Then, we subtracted the location of the peaks of $x(t)$ from the location of the peaks of z_a and averaged the result to obtain the MA in milliseconds. If z_a and $x(t)$ showed a different number of peaks, that indicated that phase-locked synchronization did not occur between the two. Fig 7 shows the result of this analysis. Each subplot shows the asynchrony between z_a and $x(t)$ for a different stimulus frequency and a pair of parameter values of λ_1 and λ_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. This analysis, however, focused on finding parameter values so that ASHLE can capture paced musician synchronization. We also want ASHLE to be able to capture musician behavior during unpaced musical performance, so carried out a third parameter analysis to understand ASHLE's behavior in the absence of a stimulus.

In our third parameter analysis we observed how, in the absence of an external

Fig 7. The asynchrony between ASHLE and a sinusoid faster or slower than ASHLE's natural frequency as a function of a narrower range of values for the frequency learning and elasticity parameters. Each cell shows the MA between ASHLE and a sinusoidal stimulus with a period 45% shorter, 30% shorter, 15% shorter, 15% longer, 30% longer, and 45% longer than ASHLE's period of natural frequency, for a pair of values for λ_1 and λ_2, which are ASHLE parameters for frequency learning and elasticity, respectively. The pair of $\lambda_1 = 4$ and $\lambda_2 = 2$ yield MA values similar to the ones that Scheurich and colleagues [15] observed in musicians synchronizing with a metronome period 30% shorter, 15% shorter, 15% longer, and 30% longer than the musician's SMP.

stimulus (i.e., $F = 0$), the parameter γ affects the return of ASHLE to f_0 when the initial values of f_p and f_a are different than f_0. This situation would occur in simulations where ASHLE learned a frequency different than its f_0 from a stimulus and then stimulus stops, or if $f_p(0)$ and $f_a(0)$ are set to initial conditions different than f_0 in the absence of a stimulus. In the experiment by Zamm and colleagues [2], musicians spontaneously started performing at a tempo faster or slower than their SMTs. A linear regression of the consecutive IBI lengths resulted in a slope that revealed that musicians progressively shrank or widened consecutive IBIs, indicating a tendency to return to their SMP. In this experiment musicians showed mean adjusted slope values in the range of -0.2 and 0.1 (see Fig 3). To simulate this behavior, we need to identify the value of γ that will result in similar slope values between consecutive IBIs. In this analysis ASHLE has an natural frequency $f_0 = 2.5$ (same as in the previous two analyses) and in different simulations the initial conditions for $f_p(0)$ and $f_a(0)$ were set to one of four different values: periof 45% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.55\times400\text{ms}}$), 30% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.70\times400\text{ms}}$), 15% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.85\times400\text{ms}}$), 15% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.15\times400\text{ms}}$), 30% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.30\times400\text{ms}}$), and 45% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.45\times400\text{ms}}$) than ASHLE's period of natural frequency. In all simulations there was no external stimulus (i.e., $F = 0$). We analyzed the behavior of ASHLE for γ values of 0.01, 0.02, 0.04, 0.08. In each simulation, ASHLE oscillated for 50 seconds with initial conditions $z_p(0) = z_a(0) = 0.001 + i0$. After each simulation, we found the location (in milliseconds) of the local maxima in the real-part of the oscillatory activity of z_a. Then we found the difference between consecutive local maxima to obtain a sequence of IBIs. Finally, we calculated the linear regression between consecutive IBIs to obtain the resulting slope. Fig 8 shows the results of this analysis. Each subplot shows the slopes obtained with different γ values and different initial conditions for $f_p(0)$ and $f_a(0)$. This analysis revealed that a value around $\gamma = 0.02$ will match the range of slope values observed in human data.

Fig 8. ASHLE slope values as a function of γ and initial frequency in the absence of a stimulus. The effect of the γ parameter on the slope values between consecutive period lengths when ASHLE oscillates without a pacing stimulus, starting at a frequency that has a period 45% shorter ($f_s = \frac{1000\text{ms}}{0.55\times400\text{ms}}$), 30% shorter ($f_s = \frac{1000\text{ms}}{0.70\times400\text{ms}}$), 15% shorter ($f_s = \frac{1000\text{ms}}{0.85\times400\text{ms}}$), 15% longer ($f_s = \frac{1000\text{ms}}{1.15\times400\text{ms}}$), 30% longer ($f_s = \frac{1000\text{ms}}{1.30\times400\text{ms}}$), and 45% longer ($f_s = \frac{1000\text{ms}}{1.45\times400\text{ms}}$) than ASHLE's period of natural frequency.

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

Behavioral data for simulation

In the task by Scheurich and colleagues [15], 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 SMT, minus the MA observed when the musician performs with a metronome that matches the SMT. Scheurich and colleagues [15] 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 SMT. 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 SMT (musician actions lagging the metronome), and negative when synchronizing with a metronome slower than the SMT (musicians actions anticipating the metronome). Additionally, the mean adjusted asynchrony grew as a function of the difference between musician SMT 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 and colleagues [15], which showed each musician's measured SMP (in their original paper, the authors mistakenly call the SMP as the 'SPR'). Using this grid over the original figure allowed us to precisely digitize each musician's SMP from the original publication. Using the musician SMP values, we simulated 20 different ASHLE models, all with the same parameter values (see model optimization below) except for the natural frequency, which has a periof that matched the SMP of a different musician. ASHLE does not generate pitch, so it only simulates the beats of the melody that musicians performed, not the melody or harmony content. This is not a problem, however, since Scheurich and colleagues only analyzed each musician's beat performance. In all simulations, initial conditions were $z_p(0) = z_a(0) = 0.001 + i0$ (indicating that ASHLE's two oscillators start with zero phase), $f_p(0) = f_a(0) = f_0$ (where f_0 is ASHLE's natural frequencyin hertz: $f_0 = \frac{1000\text{ms}}{\text{SMPms}}$. Following the procedure in the musician experiment, each ASHLE synchronized during 128 cycles with a pacing complex-valued sinusoidal stimulus $x(t) = \exp(i2\pi f_s t)$ where f_s had a period either 30% shorter ($f_s = \frac{1000\text{ms}}{0.7\times\text{SMPms}}$), 15% shorter ($f_s = \frac{1000\text{ms}}{0.85\times\text{SMPms}}$), 15% longer ($f_s = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), or 30% longer ($f_s = \frac{1000\text{ms}}{1.3\times\text{SMPms}}$) than the ASHLE's period of natural frequency. We also simulated how ASHLE would synchronize with a stimulus period 45% shorter ($f_s = \frac{1000\text{ms}}{0.55\times\text{SMPms}}$) or 45% longer ($f_s = \frac{1000\text{ms}}{1.45\times\text{SMPms}}$) 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 effector z_a 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 that Scheurich and colleagues used in their analysis, 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 natural frequency ($f_s = f_0$). We averaged the mean adjusted asynchronies observed across the 20 ASHLE models that me simulated (only different by their natural frequency value) for each stimulus period (15%, 30%, or 45% shorter or longer than ASHLE's period of natural frequency) to obtain the plot in Fig 2B. We also simulated how different ASHLE models with specific periods of natural frequency (linearly spaced between 350ms and 650ms) would carry out this task when synchronizing with a stimulus period that is 45% shorter ($f_s = \frac{1000\text{ms}}{0.55\times\text{SMPms}}$), 30% shorter ($f_s = \frac{1000\text{ms}}{0.7\times\text{SMPms}}$), 15% shorter ($f_s = \frac{1000\text{ms}}{0.85\times\text{SMPms}}$), 15% longer ($f_s = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), 30% longer ($f_s = \frac{1000\text{ms}}{1.3\times\text{SMPms}}$), or 45% longer ($f_s = \frac{1000\text{ms}}{1.45\times\text{SMPms}}$) than ASHLE's period of natural frequency. 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

Our goal was to use ASHLE to simulate the results in the musician data by Scheurich and colleagues [15]. In all our simulations in this experiment, some ASHLE parameters were always the same: $\alpha = 1$, $\beta = -1$, and $\gamma = 0.02$ (see model optimization for experiment 2 to understand how we came to this exact value for γ). The other two parameters that were always the same across simulations were $\lambda_1 = 4$ and $\lambda_2 = 2$, which we found in our parameter analysis to result in a good approximation of the musician data results by Scheurich and colleagues [15]. When we carried out the simulations in this experiment, we confirmed that the values of $\lambda_1 = 4$ and $\lambda_2 = 2$ resulted in a good fit between our simulations and the musician data by Scheurich et al. [15].

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. [2], 24 musicians individually performed a simple melody without listening to a metronome. In one control condition 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 ("Frere Jaques") four consecutive times (32 beats per repetition, 128 beats total). For each tempo, experimenters measured the IBI across each musician's entire performance and carried out a linear regression to obtain a slope, which indicates the rate of change across IOIs over the entire performance. Fig 3A shows the behavioral data with the average slope across participants in each initial tempo condition. Results showed that the slope had a tendency to be positive when performances started a spontaneous tempo faster than the SMT (IBIs becoming longer as the performances progressed), and negative when performances started at a spontaneous tempo slower than the SMT (IBIs becoming shorter as the performance progressed).

Setup, procedures and measurements

To obtain the musicians' SMP, we overlaid a grid over figure 1 from the paper by Zamm and colleagues [2], which showed each musician's measured SMP (in their original paper, the authors mistakenly call the SMP as the 'SPR'). We also overlaid a grid over figure 2

(right top panel) from the same paper by Zamm and colleagues to obtain each musician's initial rates of performance that were fast, slow, faster, and slower with respect to their SMT. Using these grids over original musician data figures allowed us to precisely recover each musician's rates of performances reported in the original study. Because our model does not generate pitch, it only simulates the beats of the melody that musicians performed, not the individual notes. Using the musicians' SMT and initial performance tempo values, we simulated 23 different ASHLE models, each oscillating with 5 different initial conditions for $f_p(0) = f_a(0)$ (matching the frequency, fast or slow compared to the SMT, and faster and slower than the SMT) (115 total simulations). We did not simulate the participant with the SMP of 665ms because its fast spontaneous tempo was signficantly faster than the rest of the particpant fast tempi, and 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 ($F = 0$). All simulations shared the same parameter values (see model optimization below) except for the natural frequency (which had a period matching the SMP of a different musician) and each simulation started with the ASHLE model having different initial conditions for $f_p(0) = f_a(0)$. After each simulation, we identified the location (in milliseconds) of local maxima in the real part of ASHLE's effector z_a. 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 $f_p(0) = f_a(0) = f_0$ was subtracted from the slope obtained in the experimental conditions ($f_p(0) = f_a(0) \neq f_0$). We averaged adjusted slopes across ASHLE models for each experimental initial frequency to obtain the plot in Fig 3B. We also simulated how different ASHLE models with specific natural frequency values (with period lengths linearly spaced between 350ms and 650ms) would carry out this task when the initial conditions for $f_p(0) = f_a(0)$ were a period 45% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.55\times\text{SMPms}}$), 30% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.70\times\text{SMPms}}$), 15% shorter ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{0.85\times\text{SMPms}}$), 15% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.15\times\text{SMPms}}$), 30% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.3\times\text{SMPms}}$), and 45% longer ($f_p(0) = f_a(0) = \frac{1000\text{ms}}{1.45\times\text{SMPms}}$) than ASHLE's period of natural frequency. 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

Our goal was to use the ASHLE model to simulate the results in the musician data by Zamm and colleagues [2]. With the exception of $F = 0$ (due to the lack of stimulus), in this second experiment we used the same parameters used in experiment 1. We wanted to use the same parameter values in order to validate our model's behavior across two different sets of behavioral results. Hence, in all simulations in this second experiment $\alpha = 1$, $\beta = -1$, and $\gamma = 0.02$. Similarly, initial conditions for Eqs (5a) and (5c) were the same between experiment 1 and experiment 2. The only set of parameters that varied between simulations were the initial conditions for $f_p(0) = f_a(0)$ and the oscillator's natural frequency (f_0).

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

Behavioral data for simulation

In another experiment, Zamm and colleagues [4] measured the SMP of 40 musicians and formed duets of musicians in two experimental groups: duets with matching SMPs

(<10ms IBI difference), and duets with 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 (64 beats total). When pairs of musicians performed the task, they first heard four metronome beats (400ms IOI) that established the common tempo. Experimenters measured the absolute asynchrony between each pair of synchronizing musicians throughout the entire performance. Because the same melody was repeated four times, they obtained a mean absolute asynchrony for each of the melody repetitions. Fig 4A shows the behavioral data reported by Zamm and colleagues (2016) with the average mean absolute asynchrony across musicians for each melody repetition and for each experimental group. Results showed that the mean absolute asynchrony between duets of musicians was larger when their SMPs did not match compared to when they matched.

Setup, procedures and measurements

To obtain the musicians' SMPs, we overlaid a grid over figure 1 from the paper by Zamm and colleagues [4], which showed each musician's measured SMP (in their original paper, the authors mistakenly call the SMP as the 'SPR'). Using the grid over the original figure allowed us to precisely recover each musician's SMP as reported in the original study. Because our model does not generate pitch, it only simulates the beats of the melody that musicians performed, not the individual notes. Using the musicians' SMP values, 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 experiment 1, except for the coupling strength between synchronizing ASHLE models (see model optimization in the next paragraph) and the natural frequency, which had a period that matched the SMP of a musician in the data by Zamm and colleagues [4]. At the beginning of the simulation, two ASHLE models were stimulated by a complex-valued sinusoid $x(t) = \exp(i2\pi f_s t)$ with an period of 400ms ($f_s = 2.5$). After these four cycles of sinusoidal stimulation, the stimulus stopped and the two ASHLE models stimulated each other with their respective z_a. That is, in each duet simulation, after four cycles of sinusoidal stimulation, the input to the ASHLE No.1 was $F_z z_{a_2}$ and the input to ASHLE No.2 was $F_z z_{a_1}$, 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 effector z_a. Then, we measured the absolute asynchrony between the two synchronizing ASHLE model's z_a, obtaining 64 absolute asynchronies for each simulation. We divided these 64 absolute asynchronies in four subsections of equal parts (16 absolute asynchronies per section) to simulate the four melody repetitions that pairs of musicians carried out in the experiment by Zamm et al. [4], and we averaged the mean absolute asynchrony in each of these four subsections. We averaged the mean absolute asynchronies across pairs of ASHLE models for each melody repetition to obtain the plot in Fig 4B. We also simulated how different ASHLE models (with natural frequency periods linearly spaced between 350ms and 650ms) would carry out this task when synchronizing for 16 cycles with another ASHLE model with an natural frequency period difference of -220ms, -110ms, -10ms, 10ms, 110ms, and 220ms. Fig 4C shows the results of these simulations, which are predictions of data that could be collected to test the accuracy of predictions made by ASHLE.

Model optimization

Our goal was to use the ASHLE model to simulate the results in the study by Zamm and colleagues [4]. With the exception of the stimulus forcing F and F_z, in this second

experiment we used the same parameters used in experiment 1 and 2. Hence, in all simulations in this experiment $\alpha = 1$, $\beta = -1$, and $\gamma = 0.02$. Similarly, initial conditions for Eqs (5a) and (5c) were always $z_p(0) = z_a(0) = 0.001 + i0$ and $f_p(0) = f_a(0) = f_0$ (matching the SMT by each musician in the duet study by Zamm and colleagues [4]). In contrast with the previous two experiments, there were two kinds of stimulation in this third experiment. During the first four cycles of each simulation, similar to experiment 1, ASHLE was simulated by a sinusoid $x(t) = \exp(i2\pi f_s t)$ with a force $F = 1$ and $f_s = 2.5$. However, 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 effector $F_z z_{a_2}$ and the input to the second ASHLE model was the first ASHLE model's effector $F_z z_{a_1}$. We found that using an $F_z = 1$ 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. We found that $F_z = 0.01$ resulted in stable synchronization, and mean absolute asynchronies matching the range of values in the behavioral data. However, we also noted that the mean absolute asynchrony was considerably smaller between pairs of ASHLE models with similari natural frequencies compared to the human 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 simulation results, in each simulation we added gaussian noise to ASHLE's effector z_a turning Eq (5c) into:

$$\frac{1}{f_a}\dot{z}_a = z_a\left(\alpha + i2\pi + \beta|z_a|^2\right) + \exp(i\angle z_p) + \mathcal{N}(\mu, \sigma^2) \tag{6}$$

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.

References

1. Ridderinkhof KR. Neurocognitive mechanisms of perception–action coordination: A review and theoretical integration. Neuroscience & Biobehavioral Reviews. 2014;46:3–29.

2. Zamm A, Wang Y, Palmer C. Musicians' natural frequencies of performance display optimal temporal stability. Journal of Biological Rhythms. 2018;33(4):432–440.

3. 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.

4. Zamm A, Wellman C, Palmer C. Endogenous rhythms influence interpersonal synchrony. Journal of Experimental Psychology: Human Perception and Performance. 2016;42(5):611.

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.

7. Repp BH. Sensorimotor synchronization: a review of the tapping literature. Psychonomic bulletin & review. 2005;12(6):969–992.

8. Repp BH, Su YH. Sensorimotor synchronization: a review of recent research (2006–2012). Psychonomic bulletin & review. 2013;20(3):403–452.

9. Mates J, Müller U, Radil T, Pöppel E. Temporal integration in sensorimotor synchronization. Journal of cognitive neuroscience. 1994;6(4):332–340.

10. 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.

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

12. Repp BH, Doggett R. Tapping to a very slow beat: a comparison of musicians and nonmusicians. Music Perception. 2007;24(4):367–376.

13. Miyake Y, Onishi Y, Poppel E. Two types of anticipation in synchronization tapping. Acta neurobiologiae experimentalis. 2004;64(3):415–426.

14. Bååth R. Estimating the distribution of sensorimotor synchronization data: A Bayesian hierarchical modeling approach. Behavior research methods. 2016;48(2):463–474.

15. Scheurich R, Zamm A, Palmer C. Tapping into rate flexibility: musical training facilitates synchronization around spontaneous production rates. Frontiers in psychology. 2018;9:458.

16. 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.

17. 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.

18. Latash M. Virtual trajectories, joint stiffness, and changes in the limb natural frequency during single-joint oscillatory movements. Neuroscience. 1992;49(1):209–220.

19. Wolpert DM. Probabilistic models in human sensorimotor control. Human movement science. 2007;26(4):511–524.

20. Stepp N, Turvey MT. On strong anticipation. Cognitive systems research. 2010;11(2):148–164.

21. 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.

22. Aschersleben G. Temporal control of movements in sensorimotor synchronization. Brain and cognition. 2002;48(1):66–79.

23. Loehr JD, Palmer C. Subdividing the beat: Auditory and motor contributions to synchronization. Music Perception. 2009;26(5):415–425.

24. Large EW, Fink P, Kelso SJ. Tracking simple and complex sequences. Psychological research. 2002;66(1):3–17.

25. Large EW, Palmer C. Perceiving temporal regularity in music. Cognitive science. 2002;26(1):1–37.

26. Kim JC, Large EW. Signal processing in periodically forced gradient frequency neural networks. Frontiers in computational neuroscience. 2015;9:152.

27. Kim JC, Large EW. Mode locking in periodically forced gradient frequency neural networks. Physical Review E. 2019;99(2):022421.

28. Righetti L, Buchli J, Ijspeert AJ. Adaptive frequency oscillators and applications. The Open Cybernetics & Systemics Journal. 2009;3(1).

29. SH S, I S. Coupled oscillators and biological synchronization. Scientific American. 1993;269(6):102–109.

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

31. Large EW, Almonte FV, Velasco MJ. A canonical model for gradient frequency neural networks. Physica D: Nonlinear Phenomena. 2010;239(12):905–911.

32. Large EW, Herrera JA, Velasco MJ. Neural networks for beat perception in musical rhythm. Frontiers in systems neuroscience. 2015;9:159.

33. 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.

34. 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.

35. 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.

36. 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.

37. 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.

38. Haken H, Kelso JS, Bunz H. A theoretical model of phase transitions in human hand movements. Biological cybernetics. 1985;51(5):347–356.

39. 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.

40. 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.

41. 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.

42. Bose A, Byrne Á, Rinzel J. A neuromechanistic model for rhythmic beat generation. PLoS computational biology. 2019;15(5):e1006450.

43. 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.

44. 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.

45. Large EW, Jones MR. The dynamics of attending: How people track time-varying events. Psychological review. 1999;106(1):119.

46. Van Der Steen MC, Keller PE. The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. Frontiers in human neuroscience. 2013;7:253.

47. Righetti L, Buchli J, Ijspeert AJ. Dynamic hebbian learning in adaptive frequency oscillators. Physica D: Nonlinear Phenomena. 2006;216(2):269–281.

48. Large EW, Kolen JF. Resonance and the perception of musical meter. Connection science. 1994;6(2-3):177–208.

49. Acebrón J, Spigler R. Adaptive frequency model for phase-frequency synchronization in large populations of globally coupled nonlinear oscillators. Physical Review Letters. 1998;81(11):2229.

50. Borisyuk R, Denham M, Hoppensteadt F, Kazanovich Y, Vinogradova O. Oscillatory model of novelty detection. Network: Computation in neural systems. 2001;12(1):1–20.

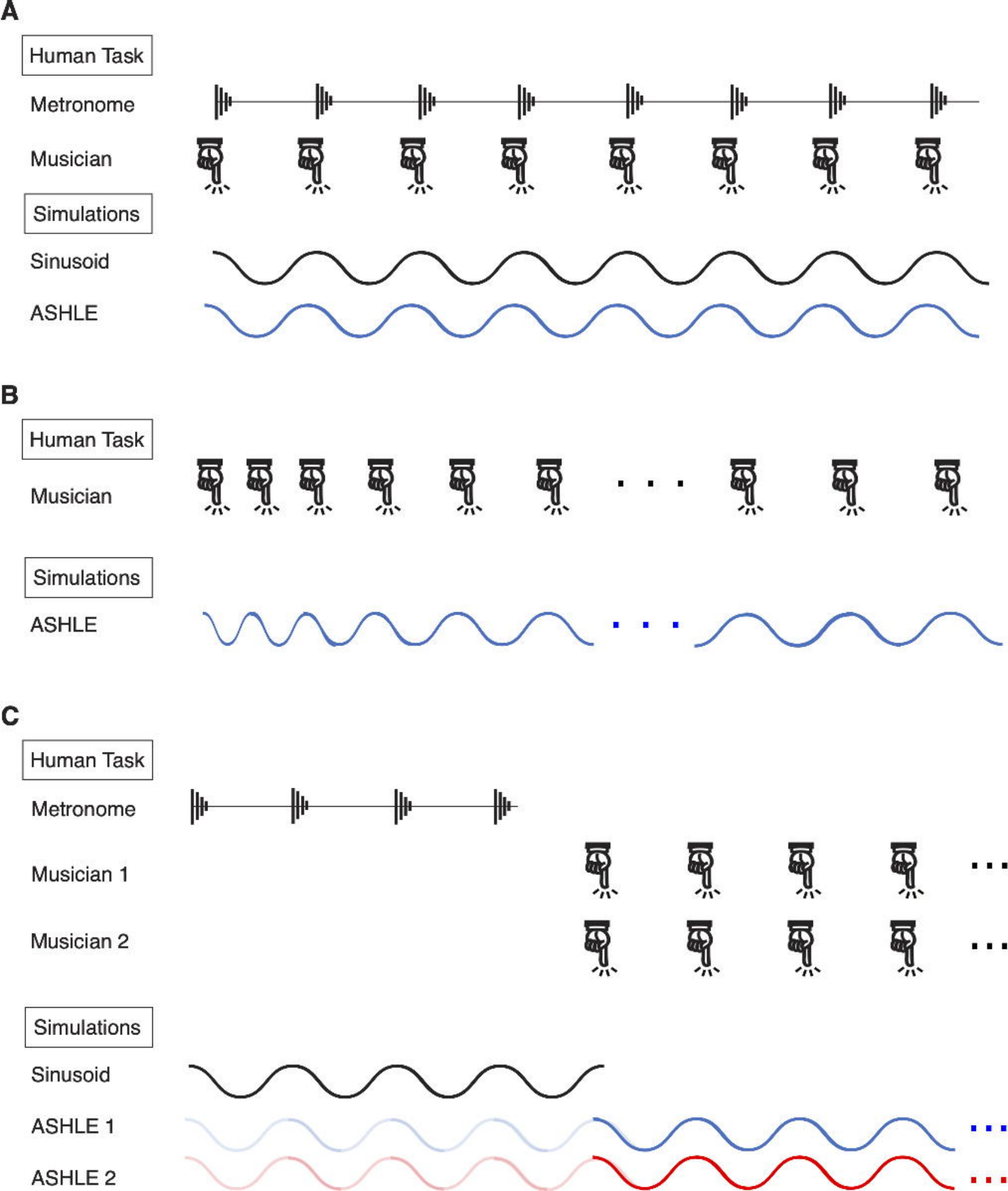
51. Ermentrout B. An adaptive model for synchrony in the firefly Pteroptyx malaccae. Journal of Mathematical Biology. 1991;29(6):571–585.

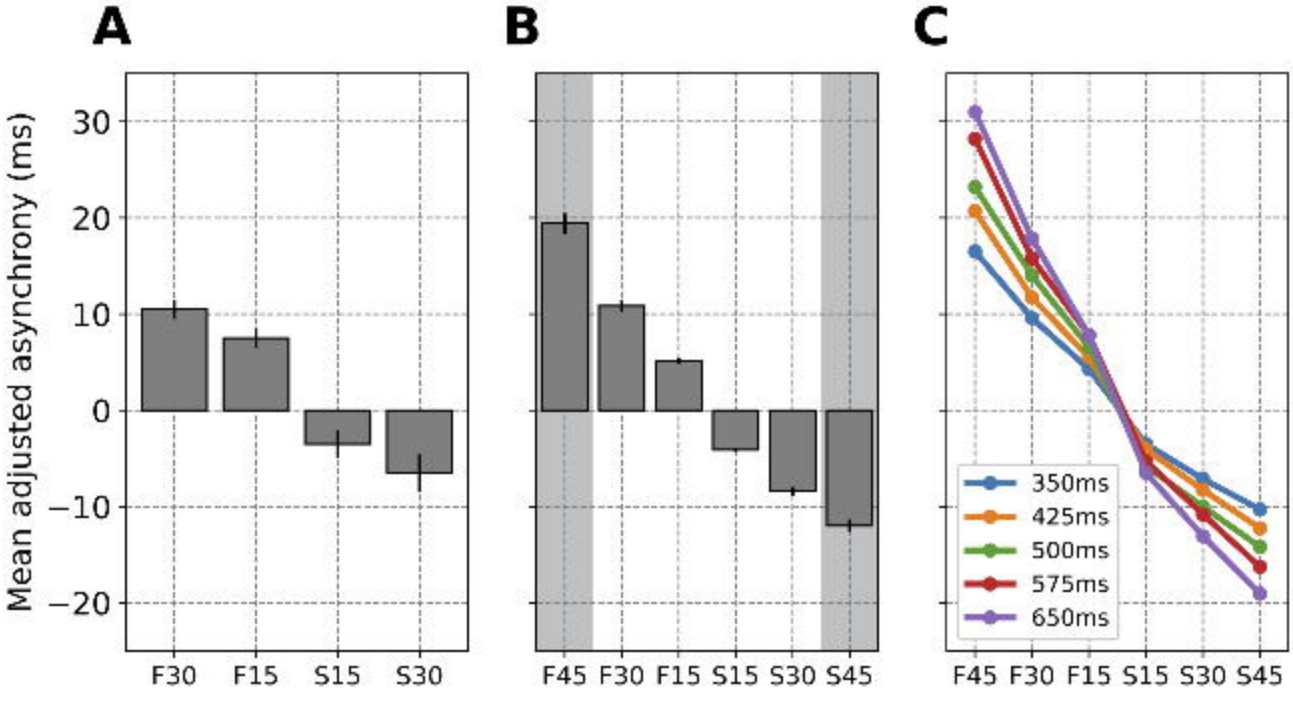
52. Nakanishi J, Morimoto J, Endo G, Cheng G, Schaal S, Kawato M. Learning from demonstration and adaptation of biped locomotion with dynamical movement primitives. In: Workshop on Robot Programming by Demonstration, IEEE/RSJ International Conference on Intelligent Robots and Systems; 2003.

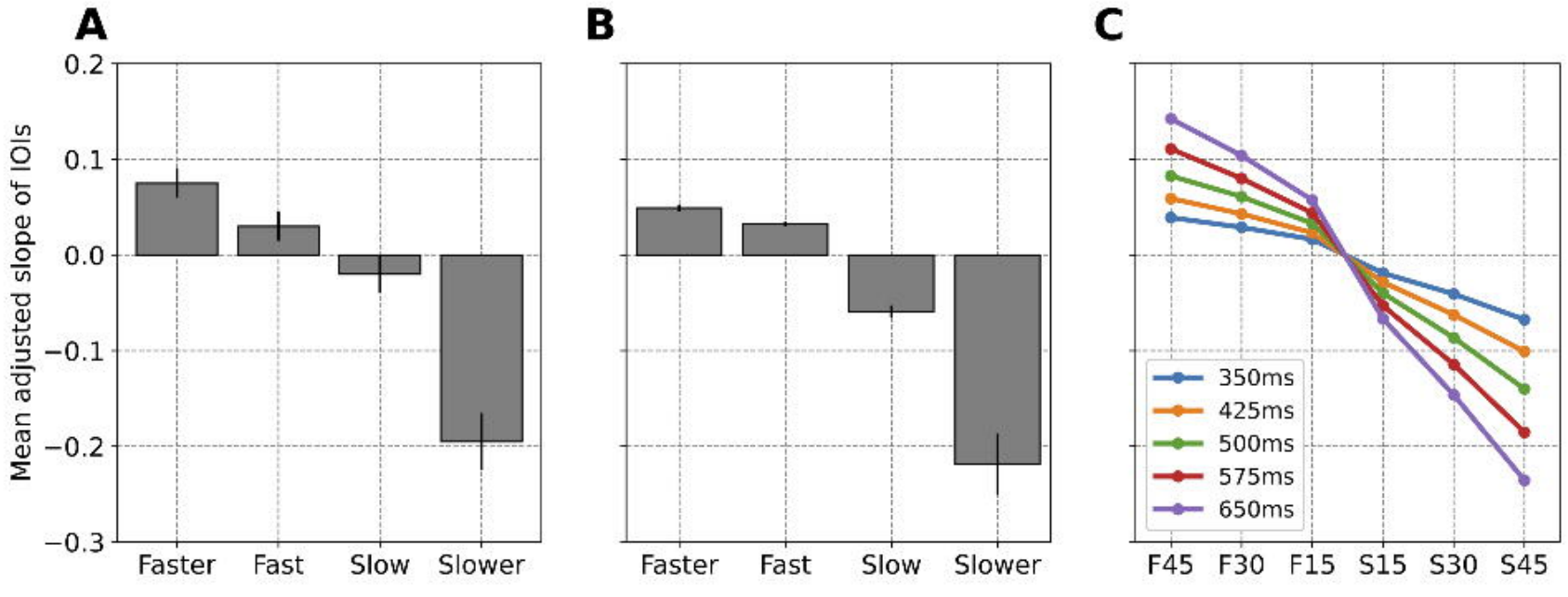
53. Nishii J. Learning model for coupled neural oscillators. Network: Computation in neural systems. 1999;10(3):213–226.

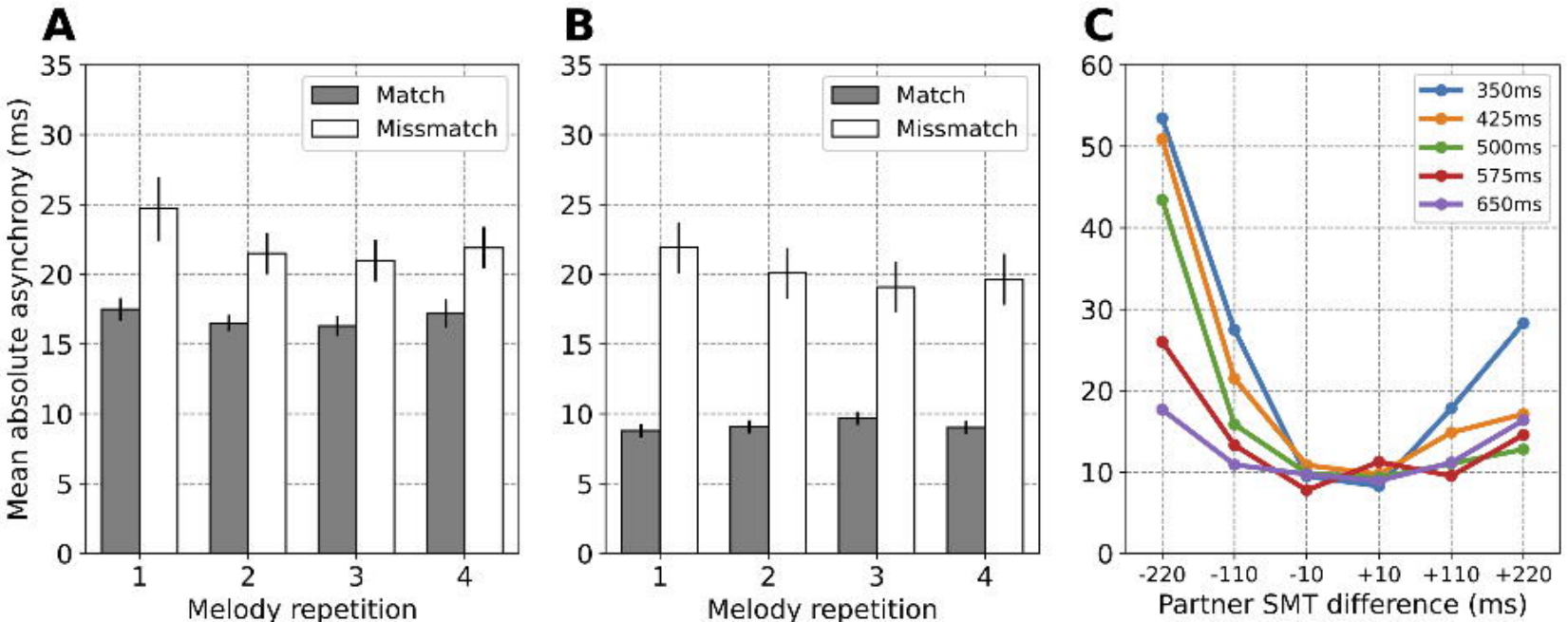
54. Lerud KD, Kim JC, Almonte FV, Carney LH, Large EW. A canonical oscillator model of cochlear dynamics. Hearing research. 2019;380:100–107.

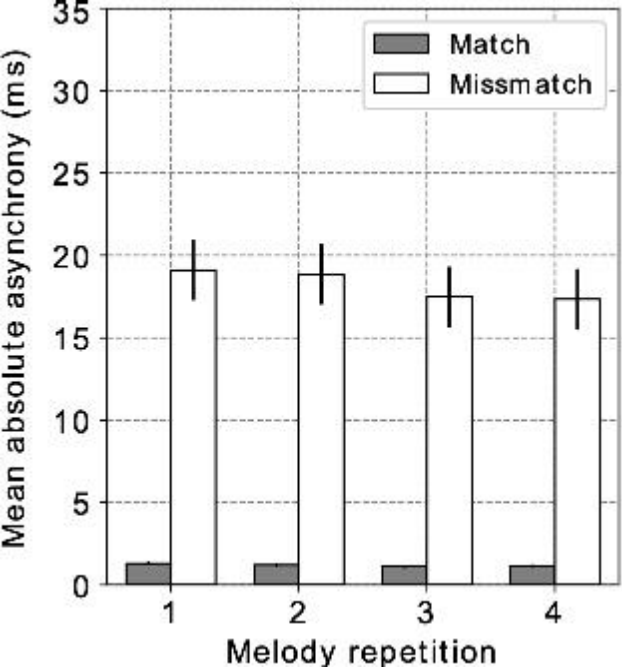
55. Tal I, Large EW, Rabinovitch E, Wei Y, Schroeder CE, Poeppel D, et al. Neural entrainment to the beat: The “missing-pulse” phenomenon. Journal of Neuroscience. 2017;37(26):6331–6341.

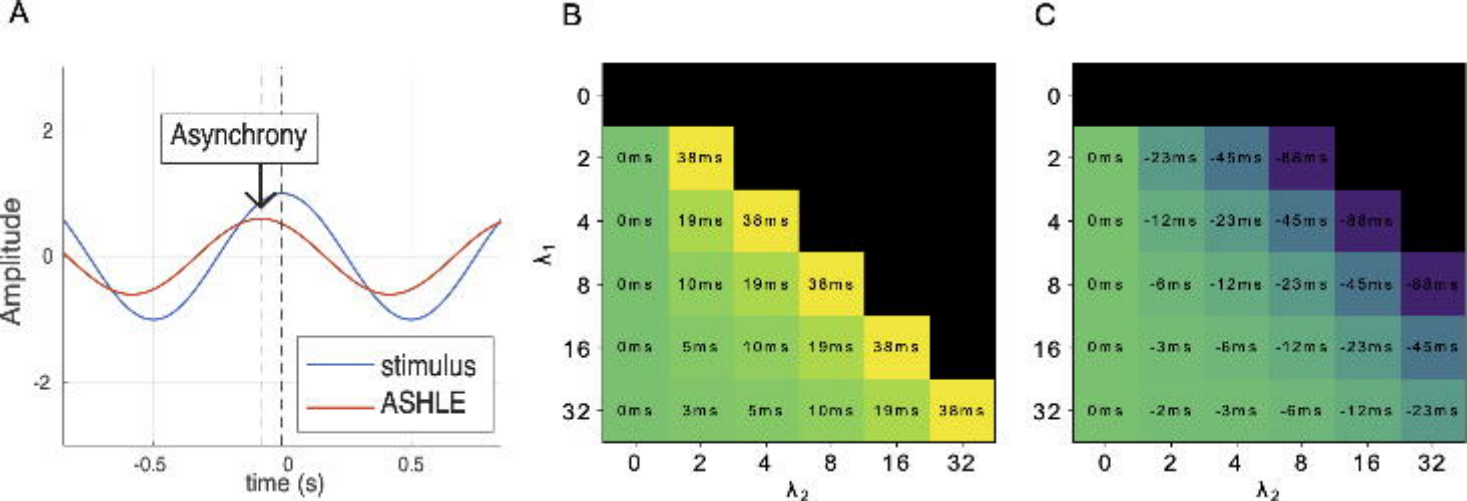

A
Human Task
Metronome
Musician
Simulations
Sinusoid
ASHLE
B
Human Task
Musician
Simulations
ASHLE
C
Human Task
Metronome
Musician 1
Musician 2
Simulations
Sinusoid
ASHLE 1
ASHLE 2

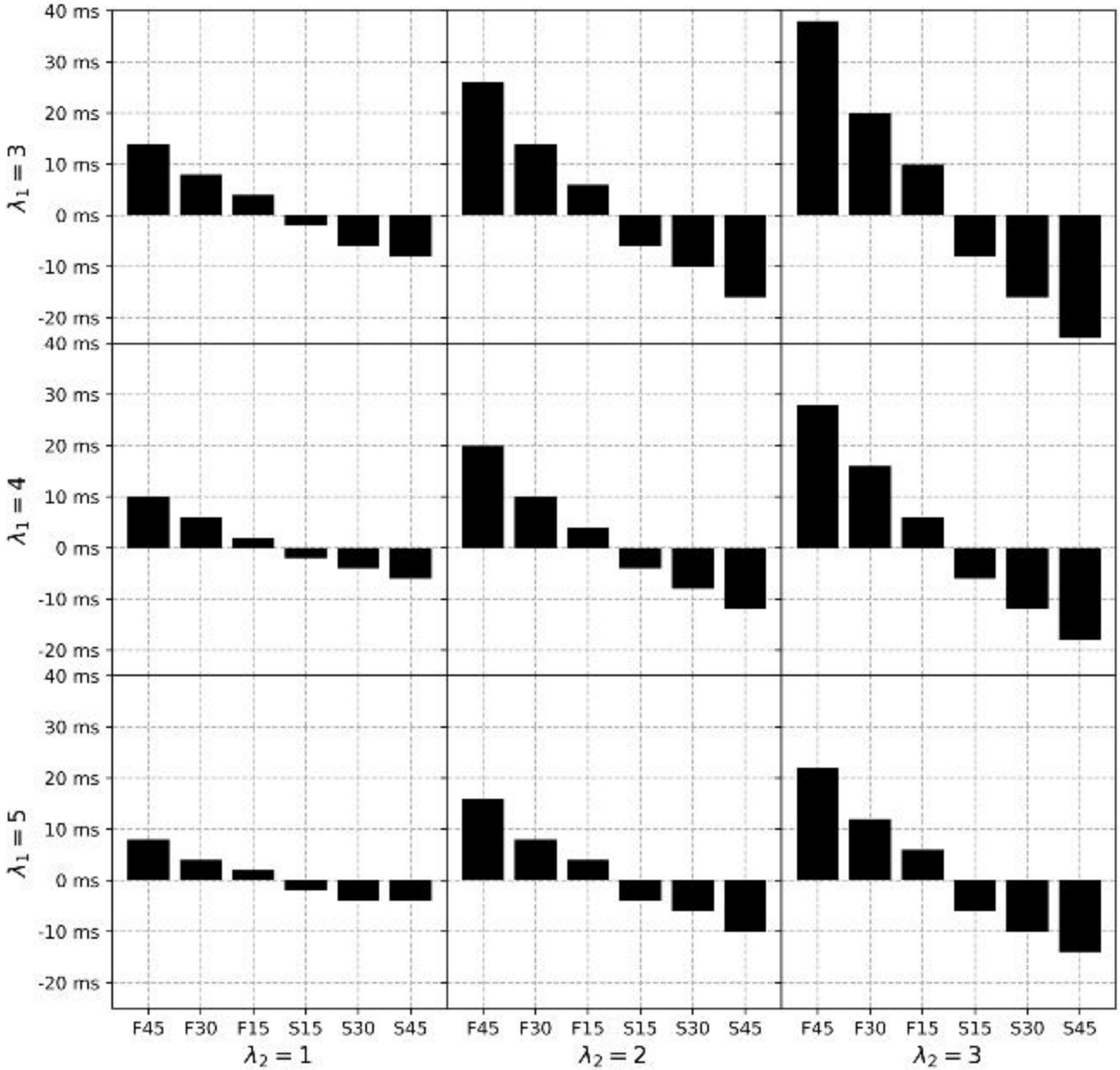

A
B
C
Mean adjusted asynchrony (ms)
30
20
10
0
−10
−20
F30
F15
S15
S30
F45
F30
F15
S15
S30
S45
F45
F30
F15
S15
S30
S45
350ms
425ms
500ms
575ms
650ms

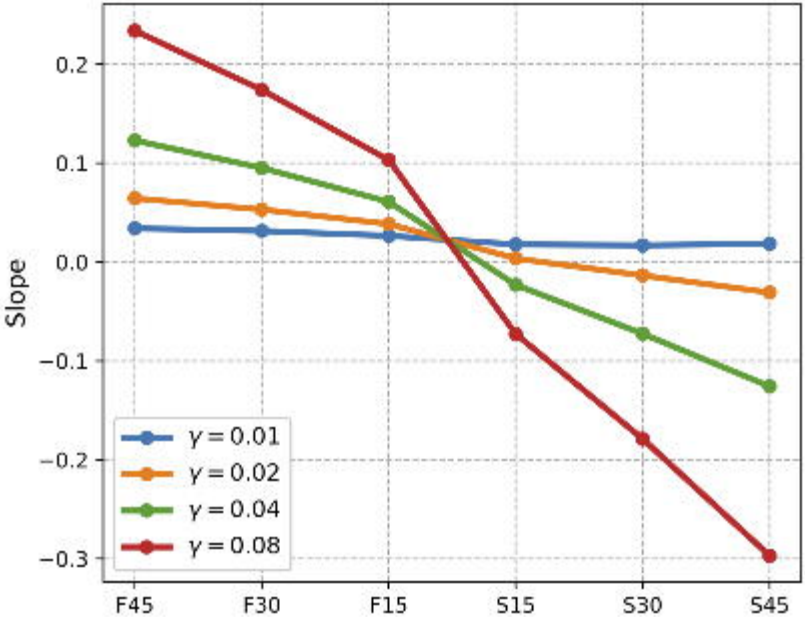

A
B
C
Mean adjusted slope of IOIs
0.2
0.1
0.0
−0.1
−0.2
−0.3
Faster
Fast
Slow
Slower
Faster
Fast
Slow
Slower
350ms
425ms
500ms
575ms
650ms
F45
F30
F15
S15
S30
S45


A
B
C
Match
Missmatch
Mean absolute asynchrony (ms)
Melody repetition
350ms
425ms
500ms
575ms
650ms
Partner SMT difference (ms)
-220
-110
-10
+10
+110
+220


Match
Missmatch
Mean absolute asynchrony (ms)
35
30
25
20
15
10
5
0
1
2
3
4
Melody repetition


A
Asynchrony
Amplitude
2
0
-2
stimulus
ASHLE
-0.5
0
0.5
time (s)
B
λ_1
0
2
4
8
16
32
0ms 38ms
0ms 19ms 38ms
0ms 10ms 19ms 38ms
0ms 5ms 10ms 19ms 38ms
0ms 3ms 5ms 10ms 19ms 38ms
λ_2
C
0ms -23ms -45ms -88ms
0ms -12ms -23ms -45ms -88ms
0ms -6ms -12ms -23ms -45ms -88ms
0ms -3ms -6ms -12ms -23ms -45ms
0ms -2ms -3ms -6ms -12ms -23ms


40 ms
30 ms
20 ms
10 ms
0 ms
-10 ms
-20 ms
$\lambda_1 = 3$
$\lambda_1 = 4$
$\lambda_1 = 5$
F45 F30 F15 S15 S30 S45
$\lambda_2 = 1$
$\lambda_2 = 2$
$\lambda_2 = 3$


Slope
0.2
0.1
0.0
−0.1
−0.2
−0.3
F45
F30
F15
S15
S30
S45
γ = 0.01
γ = 0.02
γ = 0.04
γ = 0.08