

# Small perturbations in a finger-tapping task reveal inherent nonlinearities of the underlying error correction mechanism

M. Luz Bavassi<sup>a,1</sup>, Enzo Tagliazucchi<sup>a,1</sup>, Rodrigo Laje<sup>a,b,\*</sup>

<sup>a</sup>*Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, R.S. Peña 352, Bernal (B1876BXD), Argentina.*

<sup>b</sup>*Department of Neurobiology, University of California, Los Angeles, Box 951761, Los Angeles, CA 90095, USA.*

---

## Abstract

Time estimation is critical for survival and control of a variety of behaviors, both in humans and other animals. Time processing in the few hundred milliseconds range, known as millisecond timing, is involved in motor control, speech generation and recognition, and sensorimotor synchronization like playing music or finger tapping to an external beat. In finger tapping, a mechanistic explanation in terms of neuronal activations of how the brain achieves average synchronization against inherent noise and perturbations in the stimulus sequence is still missing despite considerable research. In this work we show that nonlinear effects are important for the recovery of synchronization following a perturbation (a step change in stimulus period), even for perturbation magnitudes smaller than 10% of the period, which is well below the amount of perturbation needed to display other nonlinear effects like saturation. We build a mathematical model for the error correction mechanism and test its predictions, and further propose a framework that allows us to unify the description of the three common types of perturbations and all perturbation magnitudes with a single set of parameter values. While previous works have proposed that multiple mechanisms/strategies are used for correcting different perturbation conditions (based on fitting the model's parameters separately to different perturbation types and sizes), our results suggest that the synchronization behavior can be interpreted as the outcome of a single mechanism/strategy, and call for a revision of the idea of multiple strategies.

*Keywords:* synchronization, tapping, modeling, error correction, dynamical systems analysis

---

## 1. Introduction

Time perception and production in the scale of several hundreds of milliseconds, known as millisecond timing, is crucial for motor control, speech generation and recognition, playing music and dancing, and rapid

---

\*Corresponding author

*Email address:* [rlaje@unq.edu.ar](mailto:rlaje@unq.edu.ar) (Rodrigo Laje)

<sup>1</sup>Current address: Departamento de Física, Universidad de Buenos Aires, Pabellón I Ciudad Universitaria, Buenos Aires (C1428EHA), Argentina.

sequencing of cognitive operations such as updating working memory (Buhusi and Meck, 2005; Meck, 2005). However, our understanding of the basic mechanisms underlying this behavior is poor, and the representation of temporal information in the brain remains one of the most elusive concepts in neurobiology (Ivry and Spencer, 2004), particularly in this timing range. To date no strong consensus has been reached about which brain regions are involved in time measurement of short intervals and how they interact (Lewis and Miall, 2003; Beudel et al., 2009; Manto and Bastian, 2007; Del Olmo et al., 2008), or which is the neural mechanism responsible for the production of timed responses in this timing range (Buonomano and Laje, 2010). This stands in stark contrast to our rather comprehensive knowledge about temporal processing in other ranges, for example circadian timing (Panda et al., 2002).

### *1.1. Sensorimotor synchronization*

A paradigmatic aspect of millisecond timing is sensorimotor synchronization, which is the ability to entrain movement to an external metronome. Although a recent study reported this ability in a variety of non-human species (Schachner et al., 2009), animals display a very limited form of the behavior. In contrast it is particularly easy for humans to achieve average synchronization to a metronome and this forms the basis of all music and dance. One of the simplest tasks to study sensorimotor synchronization is finger tapping. In this task a subject is instructed to tap in synchrony with a periodic sequence of brief tones and the time difference between each response and its corresponding stimulus is recorded (see Figure 1). Despite its simplicity, this task helps to unveil interesting features of the underlying neural system and the error correction mechanism responsible for the average synchronization.

The first evidence of the existence of such a correction mechanism is the phenomenon of average synchronization itself; although no single response is perfectly aligned in time with the corresponding stimulus, average synchronization is very easy to achieve and maintain (see Figure 1; note the normal average tendency of anticipation, called Negative Mean Asynchrony or NMA). Without a correction mechanism, tiny synchronization errors or small differences between the interstimulus interval and the interresponse interval would rapidly accumulate and make the responses drift away from the stimuli, as it is very unlikely that the subject could set his/her interresponse interval exactly at the right value—even on average. This is most evident when the subject is instructed to keep tapping at the same pace after the metronome has been muted, what is called a continuation paradigm. The “virtual asynchronies” computed between the continuing taps and the extrapolated silent beats usually get quite large within a few taps (Repp, 2005), even for musically trained subjects. Note that this evidence for a correction mechanism does not preclude any kind of mechanism, since average synchronization could be achieved through either continuous adjustments (i.e. at every step), or intermittent control (i.e. once every several steps), or some other correction strategy (Gross et al., 2002), or even a mix of short- and long-range processes (Wagenmakers et al., 2004).

## 1.2. Models for finger tapping

The behavior of the underlying neural mechanism for finger tapping synchronization is usually interpreted in terms of a phenomenological error-correction function  $f$  that takes past events as inputs (including asynchronies, intervals and interval differences) and estimates the timing of the next response. This approach assumes that the underlying mechanism can be separated into a deterministic part (the correction function itself) and noise (due to inherent variability of time estimation, motor action, etc; see the seminal paper on clock and motor variance by Wing and Kristofferson (1973)). The form of the error correction function can then be generally stated as:

$$e_{n+1} = f(e_n, t_n, r_n, T_n, \dots) + \text{noise} \quad (1)$$

where  $e_n$  is the synchrony error at step  $n$ ,  $T_n$  is the interstimulus interval, and  $r_n$  is the interresponse interval; the function  $f$  could have a dependency on earlier steps as well ( $n - 1$ ,  $n - 2$ , etc.). The variable  $t_n$  is usually associated with the period of an internal timekeeper, a presumed entity proposed by Mates' influential dual-process error correction model for sensorimotor synchronization (Mates, 1994a,b). In the dual-process model the subject is capable of adjusting the phase and period of the internal timekeeper. The error correction function is usually chosen to be linear, and the inclusion of nonlinear terms is usually intended to reproduce the saturation effects observed in perturbation experiments at large perturbation magnitudes up to 50% of the stimulus period (Repp, 2002b)) or large synchrony errors (Engbert et al., 2002). The noise term in eq. (1) may have a complicated structure of its own (e.g. Vorberg and Schulze (2002)), and the experimental data in response to isochronous sequences can display long-range correlations (known as  $1/f$  noise; Chen et al. (1997); Wagenmakers et al. (2004)).

Related work within the internal timekeeper framework can be traced as far back as the pioneering work by Michon (1967), who proposed a linear predictive model that estimates the next interresponse interval based on the preceding two interstimulus intervals. Hary and Moore (1987a) proposed the influential hypothesis that the subject estimates the timing of the next tap based on either the preceding stimulus occurrence  $S_n$  or the preceding tap  $R_n$  randomly, a strategy called "mixed phase resetting" that was later shown to be equivalent to the (now usual) assumption that the estimation is based on the synchrony error  $e_n$  (Schulze, 1992). It is still under debate which perceptual information is relevant to the error correction mechanism; Schulze et al. (2005) proposed an alternative to Mates' model in which the internal timekeeper mean period is updated by the preceding synchrony error  $e_n$ , instead of the difference between the preceding interstimulus interval and timekeeper period. Even if considering only one source of perceptual information, Pressing and Jolley-Rogers (1997) showed that the subject's response to an isochronous sequence can depend on the last error  $e_n$  or the last two errors  $e_n$  and  $e_{n-1}$ , depending on whether the sequence is slow or fast, and accordingly proposed a second-order autoregressive model (Pressing and Jolley-Rogers, 1997; Pressing, 1998).

An alternative theoretical framework to explain the coordination of a subject's response to a periodic sequence is based on the concepts of self-sustained attentional oscillation, phase entrainment, and period adaptation (Large and Jones, 1999; Large, 2000). It assumes that the external rhythmic signal evokes intrinsic neural attentional oscillations that entrain to the periodicities of the sequence (Loehr et al., 2011), which is represented by a sine circle map—a nonlinear system that can show more complex entrainments than the pure 1:1 synchronization, like 2:1, etc. Although the oscillator and timekeeper models seem very different at first sight, linear timekeeper models can be seen as simplifications of nonlinear oscillator models and actually yield similar predictions for certain types of behavior (Loehr et al., 2011). Oscillator models probably surpass timekeeper models when stimuli with multiple periodicities are considered (Loehr et al., 2011); to the best of our knowledge, however, the extent to which they accurately reproduce the transient behavior when abrupt perturbations are considered has yet to be demonstrated—the perturbations considered by Loehr et al. (2011) were not abrupt step changes, but slow linear increases or decreases in tempo, which in fact can be seen as quasi-stationary if the model is in a high dissipation regime.

### *1.3. Nonlinear behavior*

Several reports bore a few hallmarks of nonlinearity in finger tapping tasks. Repp (2002b) showed that the response to a phase-shift perturbation displays at least two distinctive nonlinear features: asymmetry and saturation in the phase correction response function (PCR, equal to the difference between the time of occurrence of the first tap after the perturbation and the time when this tap would have been in the absence of perturbation). Asymmetry was evident as smaller values of the PCR for negative than for positive shifts, but only for perturbation magnitudes greater than  $\pm 50$  ms ( $\pm 10\%$ ). Saturation effects, displayed as a shallower slope of the PCR function for large perturbations than for small perturbations, were naturally more evident at still greater perturbation magnitudes. Interestingly, the author also found an asymmetry in the standard deviation of the PCR: the variability was higher for positive than for negative perturbations, but again only outside the range  $\pm 50$  ms. In a very recent work, Repp (2011) reported similar asymmetries in response to phase-shift perturbations. The PCR to positive phase shifts was again larger than the PCR to negative phase shifts. A significant asymmetry was only found for large perturbations (greater than  $\pm 75$  ms), and led the author to the interpretation that it is easier to delay than to advance a tap in response to a perturbation.

So far, asymmetries have been largely reported for phase shifts only, and large perturbation sizes or large asynchronies only. It is remarkable that these findings have not yet been reflected in substantial changes to the models—other than considerations regarding the saturation of the response and only in an isochronous task, like the work by Engbert et al. (2002) who proposed a nonlinear model to explain a small subset of the data in an isochronous finger tapping task. They showed that an error correction model with a saturating function  $\tanh(e_n)$  was consistent with the experimental results, and interpreted it as a saturation in the

subject’s response when the synchrony error  $e_n$  is large.

The nonlinear behavior, however, seems pretty robust, and sooner or later the models will have to address it (Repp, 2011). The absence of nonlinear features in the models can probably be explained because the linear approximation appears to be still valid for small perturbation sizes. However, in this work we demonstrate asymmetric responses to step-change perturbations for perturbation magnitudes smaller than 50 ms (10%), and accordingly propose a nonlinear (quadratic) model. There is an important *qualitative difference* between asymmetric data and a linear model—after symmetric perturbations, a linear model can only yield symmetric responses (Loehr et al., 2011). This difference cannot be address quantitatively with a better fitting of a linear model, and calls for a revision of the validity of linear models within the usual range of small perturbation magnitudes.

#### 1.4. *Different strategies for different conditions*

Perturbation experiments are usually performed to probe the response of the system, most notably in the form of either a “step change”, or a “phase shift”, or an “event-onset shift” (see Figure 2), where both the magnitude of the perturbation and the time of occurrence are unexpected. A surprisingly common approach in the field—regardless of the chosen framework, whether timekeeper- or oscillator-based—is to fit the model’s parameter values separately to different conditions, thus yielding different parameter estimates for different perturbation types and even for different perturbation magnitudes within the same perturbation type (see e.g. Thaut et al. (1998); Repp (2001b); Schulze et al. (2005); Large et al. (2002); however see a very recent exception to this common choice by Loehr et al. (2011)). Then although the parameter names are the same within each model, it is allowed for instance that the coefficient of the period correction be small for some perturbations and large for some other perturbations, effectively changing the model’s correction strategy and thus the interpretation of the data. For instance, a study that explicitly suggested separate strategies for correcting large and small step-change perturbations is that of Thaut et al. (1998). The authors found a significant difference in their model’s fitted parameter values for large and small step-change perturbations, including a huge difference of two orders of magnitude between the extreme values of the fitted  $\beta$  (period correction), ranging from  $-0.496$  to  $-0.006$ , effectively making the proposed period correction term disappear for some perturbations. This, together with the observation of apparently qualitatively and quantitatively different time evolutions for the experimental series, led the authors to state, “the observed multiple synchronization strategies are expressed in our brief mathematical model through adjustments in the equation constants”.

Within the linear, dual-process model of Mates (1994a,b), the proposed period correction process was shown to be dependent on the subject’s awareness of a tempo perturbation (Repp and Keller, 2004). The authors fitted the model separately per tempo condition and found a significative difference between the values of the period correction coefficient  $\beta$  contingent on the detection responses (although they acknowledged

that the model fits ranged from “excellent to very poor”, that “the two-process model of error correction did not fit the data as well as [they] would have liked”; they did not showed either any direct comparison between experimental and model time series). Awareness of the perturbation is undoubtedly an important factor of the underlying mechanism responsible for the synchronization behavior. However, the fact that a compound model (period + phase correction) can reproduce the observed behavior only after separate fitting is arguably weak evidence supporting any part of the model, as discussed in the following.

An important shortcoming of this usual practice is that it does not offer any explanation as to how the subject would “choose” the correct strategy at the beginning of the trial, since he/she is unaware of the magnitude of the upcoming perturbation and thus it would be impossible to “shift gears” in advance. Even in the hypothetical case of the subject being able to immediately choose the correct strategy right after the first perturbed step, a second timing mechanism is needed in addition to the usual error correction to make the choice of strategy. Therefore, the procedure of separate fitting implicitly and necessarily assumes that there should be an additional mechanism for quickly selecting the appropriate set of parameter values, i.e. selecting different correction strategies described by different parameter values. Indeed, Schulze et al. (2005) estimated their models parameters separately both per tempo and tempo change condition and noted, “there must exist additional control mechanisms that determine when the period adjustment mechanism is started and stopped (e.g., by setting the period correction gain)”. Thus a fundamental question remains, can we describe the behavior with a simple model, a single set of parameter values, and no additional hypotheses? Can all these various, seemingly different observed responses be part of a broader spectrum of possible responses of a single mechanism/strategy?

### 1.5. *A single mechanism/strategy*

The sensorimotor synchronization behavior is likely to draw on several distinct neural processes, namely time perception, interval comparison, error detection, time production, and motor execution (Repp, 2005). The question posed above relates to whether this likely superposition of neural processes leads to different strategies for different perturbations plus an additional mechanism for choosing among strategies, or whether the whole behavior can be interpreted as the result of a single mechanism/strategy represented by a relatively simple model with a fixed number of terms and a fixed, single set of parameter values.

In this work we search for both theoretical and experimental evidence supporting the hypothesis of a single underlying mechanism/strategy for the sensorimotor synchronization behavior in humans. By this we mean that, although several distinct neural processes are probably involved as we pointed out above, the error-correction mechanism resulting from the interplay of such processes can be interpreted as a single entity, as opposed for instance to separate strategies for correcting perturbations of different signs or magnitudes. Based on the dynamical constraints that the observed behavior poses on the possible mechanisms, we propose a mathematical model for the error correction function  $f$  (eq. 1), without assuming the existence of

an internal clock of any particular kind, or any other hypothesis on the actual neural mechanism in charge of achieving average synchronization. We search for a unified framework in which the model accounts for the three most common types of perturbation experiments and all studied perturbation magnitudes with a single set of parameter values. Although not a proof, the success of such a unification effort would be suggestive evidence for the oneness of the underlying mechanism.

## 2. Theoretical Results

### 2.1. General considerations

We choose the observable synchronization error  $e_n$  as our main variable because of its fundamental nature (Chen et al., 1997), and propose an error correction model in the form of a map like eq. 1—that is, we propose the shape of the function  $f$ . We assume as our working hypotheses that it is possible to identify a deterministic component within the general mechanism of error correction (i.e. we assume the usual separation between a deterministic rule and noise), and that the qualitative behavior of the deterministic component can be described in terms of a small number of variables (i.e. we are not proposing a realistic neural model, but a phenomenological behavioral model). This separation between a deterministic component and inherent noise is particularly important when studying perturbations, where the deterministic rule is likely to have more importance than in an isochronous setting (which is the usual setting for studying the noise component).

One of the most influential models for the error correction mechanism (Mates, 1994a,b) included motor and perceptual delays, and a distinction between external (observable) and internal (psychological/neural) variables. At this point we do not assume anything about the psychological or neural basis of the mechanism, and instead search for the dynamical constraints that the observed behavior sets on the possible underlying mechanisms.

### 2.2. Theoretical implications of previous experimental studies

Several dynamical constraints on the possible models for the underlying error-correction mechanism become evident after careful reviewing of the literature on finger-tapping tasks. In the following we pinpoint the most important findings and interpret them theoretically in order to build our model.

#### 2.2.1. Thresholds are not needed

Perturbations to the stimulus sequence in an otherwise isochronous finger-tapping task are very informative for probing the mechanism of synchronization. One of the most common perturbations is a constant shift of all tones from a certain point in the sequence, called a “phase shift”. Indeed, phase-shift perturbations were used to show that even subliminal changes smaller than 4% of the stimulus period (the approximate threshold for perceptual detection of interval changes in musicians) led to a correction behavior (Repp, 2000). Repp showed that in all cases, irrespective of perturbation magnitude or sign, the correction mechanism

was engaged already at the first tap after the (unexpected) perturbation and the return to the baseline was monotonic (see Figure 3A).

Although recent work has shown an asymmetric response to positive and negative phase shifts (Repp, 2011), and a nonlinear response for large perturbation magnitudes (Repp, 2002b), it is clear that the qualitative shape of the time evolution of the response is always a monotonic recovery of average synchronization, at least within the range of perturbation magnitudes studied here. In addition, error correction is engaged even at the smallest perturbation magnitude studied—smaller than the perceptual detection threshold for auditory-auditory or auditory-kinesthetic temporal order (about 20 ms; Hirsh (1959); Hirsh and Scherrick (1961); Repp (2000)), or phase-shifts (at least 10 ms; Repp (2000)), or interval perturbations (about 4% for time intervals; Drake and Botte (1993); about 2% for tempo changes; McAuley and Kidd (1998); see Repp (2000) for more references therein). Several other studies have also shown that subjects respond to subliminal perturbations (Hary and Moore, 1987b; Madison and Merker, 2004). This suggests that the processes involved in sensorimotor synchronization have access to more accurate timing information than conscious processes of temporal judgment (Repp, 2000). An alternative, more parsimonious interpretation is based on the hypothesis that the subject times his/her next tap, in part, with reference to the shifted tone and, in part, with reference to the previous tap (Repp, 2002a, 2005). In any case, these results imply that there is no need to incorporate detection thresholds into the models.

### 2.2.2. *Isochronous sequences and dimensionality*

Even the synchronization to an isochronous sequence, i.e. a periodic sequence without perturbations, can show nontrivial features. In such studies, the period of the sequence is varied among trials, with values typically between 200 and 750 ms (although shorter and longer periods are also tested). Two independent works have shown that the correction mechanism is fed by the last synchronization error  $e_n$  for slow sequences, but by the last two synchronization errors  $e_n$  and  $e_{n-1}$  for fast sequences (Pressing and Jolley-Rogers, 1997; Semjen et al., 2000). That is, the experimental data was best fitted by

$$e_{n+1} = \alpha e_n \tag{2}$$

for slow sequences, and by

$$e_{n+1} = \alpha e_n + \beta e_{n-1} \tag{3}$$

for fast sequences. Pressing and Jolley-Rogers (1997) interpreted this result based on a number of observations, most notably that the fastest periods are less than or of the order of the auditory reaction time (Luce, 1986), as well as typical estimates for decision-based error correction times (Gibbs, 1965). Some degree of overlap between the processing of consecutive perceptual information is then expected in this condition. A different, likely interpretation is based on hierarchical metrical structure, with the beat moving to a higher (i.e. slower) level for faster sequences (Repp, 2008).

We model this result as follows. Assume that the correction function is linear, at least for non-perturbed sequences. We choose to represent our system with a two-variable model fed by only the last step, instead of a single-variable model fed by the last two steps. This is possible because there is a correspondence between the two representations; a simple algebraic manipulation leads to the following identification:

$$\begin{pmatrix} e_{n+1} \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e_n \\ x_n \end{pmatrix} \Rightarrow e_{n+1} = Ae_n + Be_{n-1} \quad (4)$$

for appropriately chosen parameter values  $\mathbf{M} = (a, b; c, d)$  such that  $\text{tr}(\mathbf{M}) = A$  and  $\det(\mathbf{M}) = -B$ . Here the variable  $x_n$  is an auxiliary variable of dynamical origin and in principle is not intended to represent any particular biological/neural/perceptual/physical entity (see Discussion). We interpret the crossover between slow and fast sequences (eqs. 2 and 3) as a reduction in the effective dimensionality of the 2D system, since setting  $A \neq 0$  and  $B = 0$  in the 1D system is equivalent to setting  $\text{tr}(\mathbf{M}) \neq 0$  and  $\det(\mathbf{M}) = 0$  in the 2D system, which can be achieved by setting one of the eigenvalues to a value much smaller than the other. Indeed, if  $\lambda_2 \ll \lambda_1$  then  $\text{tr}(\mathbf{M}) = \lambda_1 + \lambda_2 \approx \lambda_1$  and  $\det(\mathbf{M}) = \lambda_1\lambda_2 \ll \lambda_1$ .

We have two main reasons to choose a 2D representation instead of a 1D, two-step model. First, note that the correspondence described by eq. (4) is strictly valid for linear models only—it breaks down as soon as we include a nonlinear term. This is due to the fact that the algebraic manipulation needed to switch between the two representations involves solving for some of the variables, i.e. finding an inverse function; and the inverse of a function is not always a function, which is readily evident when quadratic terms are included (for instance the inverse of  $z_n = e_n^2$  would be the multivalued curve  $e_n = \pm\sqrt{z_n}$ ). Second, our choice allows us to use the tools of Nonlinear Dynamics theory and the powerful, geometrical interpretation of the phase space, which is particularly helpful to model the two significant features we point out—overshoot and asymmetry.

### 2.2.3. Perturbations and dimensionality

A second common perturbation in the finger tapping literature is a sudden increase or decrease in the period of an otherwise isochronous sequence, what is called a “step change”. A most interesting finding is evident in the experimental time series displayed in Figure 3B: the asynchronies in response to the step change of 10% of the stimulus period exhibit considerable overshoot before approaching the new baseline (Thaut et al., 1998; Michon, 1967). This result is most important regarding the building of our model. The overshoot after a perturbation cannot be reproduced by a one-dimensional (1D) map fed by only the last step, since that would violate the deterministic nature of the equation—no value of the variable can have two different possible futures (Schöner, 2002). The only behavior that such a model could display after a perturbation is either a monotonic decay to the baseline value or a monotonic divergence from it, whatever nonlinear terms it may have. In order to reproduce an overshoot, a two-dimensional (2D) map is needed

instead—or alternatively a 1D model with a two-step dependence (left and right sides of eq. 4, respectively). As before, we choose a 2D representation.

#### 2.2.4. *Asymmetry*

Another most interesting finding from the step-change perturbations is that the overshoot is only displayed for positive perturbations, i.e. when the period of the sequence is increased but not when it is decreased (Thaut et al., 1998). This asymmetric behavior can be readily observed in the time series in Figure 3B: symmetric negative and positive step-change perturbations have an asymmetric effect on the systems response. After a step decrease in the interstimulus interval (a negative perturbation) the subject monotonically approaches the baseline, but after a step increase (a positive perturbation) the baseline is only approached after overshooting. It can be argued that this observation is evidence of two distinct underlying mechanisms, or two fundamentally different ways of perturbing the system (see Discussion); we assume, however, that the asymmetric behavior is inherent to a unique underlying system, and thus incorporate it in the model by introducing appropriate quadratic terms to break the symmetry displayed by the linear part.

Note that the asymmetry discussed above appears after a step-change perturbation of  $\Delta = \pm 50$  ms ( $\pm 10\%$ ), that is just within the usual range where linear behavior is assumed. Repp (2011) also reported asymmetries, but in response to phase-shift perturbations and for perturbation magnitudes greater than  $\pm 75$  ms (with a pre-perturbation interstimulus interval of 500 ms).

#### 2.2.5. *Saturation effects*

Finally, when the perturbations are large (up to 50% of the basal interstimulus interval), saturation effects appear. Repp (2002b) applied phase shifts of variable magnitude and both signs to the stimuli sequence and recorded the relative shift of the tap immediately following the perturbation (known as the phase correction response or PCR). He found that the amount of correction was a linear function of the perturbation magnitude for small perturbations (between 10% and -10% of the initial stimulus period), but the slope of the function decreased for larger perturbations, giving the PCR function a sigmoid-like shape possibly reflecting a saturation in the correction ability (see Figure 3C). Also Engbert et al. (2002) found evidence of saturation effects in isochronous sequences. They tested the synchronization time series for unstable periodic orbits and rejected the null hypothesis of an underlying linear system in about 4% of the series in the experimental dataset, and showed that a sigmoid-shaped correction function is consistent with that finding. Engbert et al. (2002) assumed that the subject’s ability for correcting deviations is different for small and large deviations—a common, biologically plausible assumption—and modeled this saturation effect by adopting a correction function of a sigmoid shape ( $\tanh$ ). This implies an asymptote that would probably be only evident in the extreme case (if any) of a very large perturbation, as the ones in Repp (2002b). For smaller perturbation magnitudes, the same effect is seen as a decreasing slope in the PCR function. In a low-order modeling approach as ours, the existence of an asymptote or saturation is not

represented by a full tanh function, but by its next-to-linear term in the Taylor series, which is a cubic term (consideration of even larger perturbation sizes would probably need the inclusion of a fifth-order term, etc.). Since we focus on small perturbations (10% or less) to limit the amount of nonlinear phenomena in the data, for the sake of simplicity—and to avoid the likely masking of the smaller, second-order effects—we choose to disregard saturation effects and thus we do not incorporate such nonlinearity into our model.

### 2.3. Model equations

According to our previous considerations, we propose the following two-dimensional model for the correction function in a finger tapping task:

$$\begin{aligned} e_{n+1} &= ae_n + b(x_n - T_n) + \alpha e_n(x_n - T_n) + \beta e_n^2 + \gamma(x_n - T_n)^2 \\ x_{n+1} &= ce_n + d(x_n - T_n) + x_n \end{aligned} \tag{5}$$

where  $e_n$  is the (observable) synchronization error, or asynchrony, at step  $n$ ;  $x_n$  is an auxiliary variable of dynamical origin, not intended in principle to represent any biological/neural/perceptual/physical entity but necessary to reproduce the overshoot and the 2D behavior;  $T_n$  is the interstimulus interval at step  $n$ , entering as a parameter;  $a, b, c, d$  are the linear coefficients; and  $\alpha, \beta, \gamma$  are the coefficients of the nonlinear terms, needed to reproduce the asymmetric response to step-change perturbations.

It is easy to see that for isochronous sequences, where  $T_n = T$  for all  $n$ , this model has a fixed point at  $e_n = 0$  and  $x_n = T$ . The values for the linear coefficients should be such that the fixed point is stable—that is, the asynchrony  $e_n$  tends to zero and  $x_n$  tends to the value of the constant interstimulus interval  $T$ —, and in particular a stable node. We discard stable spirals because the experimental data does not show any consistent oscillation while approaching synchrony, and we discard stars and degenerate nodes because of the need for carefully tuning the parameters (the linear coefficients should be set such that the two eigenvalues are exactly equal). For a view of the most relevant trajectories in our model’s phase space, see Figure 4. The set of parameter values used throughout this work is displayed in Table 1.

Note that we do not account for the constant NMA generally displayed by the subjects, because the target of our modeling effort is how the system responds to perturbations—the transient, not the equilibrium—, and so we shift the pre-perturbation baseline (the point of subjective synchrony) to zero by simply setting  $e_n \rightarrow e_n - \text{NMA}$  (see Materials and Methods). Other authors have extended their models such that they accommodate without explaining nonzero mean asynchronies (Thaut et al., 1998; Schulze et al., 2005), or have included terms representing generic perceptual delays that effectively work as free parameters to accommodate a constant shift (Semjen et al., 1998). One successful account for NMA is based on the difference between somatosensory and auditory feedback information processing speeds in the brain (see Repp (2005) for a review). Models for sensorimotor synchronization like ours (e.g. Thaut et al. (1998); Schulze et al. (2005); Repp (2001b)) do not take into account a detailed description of the processing mechanisms of sensory

information—they propose a phenomenological approach to the error correction function, due to the multitude of neural substrates likely involved. In fact, the only mathematical models that can address the NMA in synchronization tapping—without actually modeling the details of the processing mechanisms—are coupled oscillator models or circle maps (Large, 2000; Loehr et al., 2011), where the NMA arises as the stationary phase difference between the oscillatory attentional focus (“forced oscillator”) and the periodic metronome (“forcing stimulus”). The phase difference depends on the detuning between the natural frequency of the oscillator and the frequency of the stimuli. In summary, we emphasize that few phenomenological models account for the NMA; none of them accounts for it with a single set of parameter values; none of them accounts for it under perturbations; and none of them account for the pre-post perturbation shift in NMA as a function of the perturbation size (which is a novel finding described below).

### 2.3.1. Perturbations in the model

A most important feature of this model is how it responds to perturbations. First, we assume that the interstimulus interval enters the model as a parameter  $T_n$ , and thus we interpret a step-change perturbation as a single, step change in the value of  $T_n$  at step  $n = 0$  (see Figure 2A):

$$\begin{aligned} T_n &= T & (n < 0) \\ T_n &= T + \Delta & (n \geq 0) \end{aligned} \tag{6}$$

Second, a phase-shift perturbation in this framework is not an elementary perturbation but it is formed by two consecutive step changes. The interstimulus interval changes twice such that it remains the same after the perturbation ( $T \rightarrow T + \Delta \rightarrow T$ ), and the net effect is a time shift of all subsequent events (see Figure 2B). Finally, an event-onset shift perturbation is the composition of three consecutive step changes, and the final interstimulus interval is the same ( $T \rightarrow T + \Delta \rightarrow T - \Delta \rightarrow T$ , Figure 2C).

All perturbations considered in this work are unexpected, meaning that the subject knows that a perturbation is possible but does not know whether it will actually occur, or the time, or its magnitude or its sign. As the subject is unaware, this results in a forced error at the step of the perturbation, on average equal to the magnitude of the perturbation (with opposite sign). We can conceptualize this by unfolding the variable  $e_n$  into “predicted  $e_n$ ” and “actually observed  $e_n$ ” (the difference being, on average, equal to the perturbation size at the step when a perturbation occurred, or zero otherwise). In a step-change perturbation experiment, the predicted  $e_n$  and the actually observed  $e_n$  are both zero until the perturbation arrives of size  $\Delta$ . At the perturbation step the predicted  $e_n$  is still zero (because the perturbation is unexpected), but the actually observed  $e_n$  will be  $-\Delta$ . So, in order to let the model “know” that a perturbation occurred at step  $n$ , the parameter  $T_n$  should be changed to the new value  $T + \Delta$  and the value of the variable  $e_n$  should be reset and overridden by the value of “actually observed  $e_n$ ”, that is  $-\Delta$ . If the variable  $e_n$  at the perturbation step has a value that is different from zero because of previous history, then the value of “actually observed  $e_n$ ” is the opposite of the perturbation size  $-\Delta$  plus the current  $e_n$  (with the corresponding

sign), as it naturally occurs in the experiment.

Thus, in this model, any time the sequence is perturbed the variable  $e_n$  should be reset as described above to reproduce the forced error, which has the effect of resetting the initial conditions of the subsequent evolution (see Figure 4). For a sequence with original period  $T$  that is perturbed by  $\Delta$  at step  $n = 0$ , and assuming that the asynchrony before the perturbation was zero, the correct way to represent a step-change perturbation in the model is then

$$\begin{aligned} T_0 &= T + \Delta \\ e_0 &= -\Delta \end{aligned} \tag{7}$$

with  $T_{n>0} = T + \Delta$  and subsequent evolution of  $e_{n>0}$  given by the model's equations.

#### 2.4. Model predictions for step-change perturbations

The geometrical organization of the trajectories in our model's phase space (Figure 5A) allows us to make a prediction about the overshoot after a step-change perturbation. Note in Figure 5A that the approach to the fixed point after a positive perturbation (i.e. from below) always displays an overshoot in  $e_n$ , and the amount of overshoot depends on the perturbation magnitude. On the other hand, the overshoot is almost undetectable for negative perturbations (i.e. from above), no matter the perturbation magnitude.

We tested our model's prediction by performing a finger-tapping experiment with step-change perturbations. The published data for step-change perturbations deals with short trials and perturbation magnitudes of only up to 5% (Repp, 2001b), or longer trials and larger magnitudes but a limited number of intermediate magnitudes (Thaut et al., 1998). Since we needed long trials in order to let the subject reach the equilibrium possibly after an overshoot, and the behavior for intermediate perturbation magnitudes was important to demonstrate the goodness of the model, we performed our own experiment with 60-tone trials and step-change perturbations of  $\pm 50$ ,  $\pm 40$ ,  $\pm 30$ ,  $\pm 20$ , and  $\pm 10$  ms (that is, between  $\pm 10\%$  and  $\pm 2\%$ , see Materials and Methods). The model was fitted with a genetic algorithm to the  $\pm 50$  ms time series only (Figure 5B), i.e. the extreme perturbations only.

### 3. Materials and Methods

#### 3.1. Participants

The experiment consisted in a finger-tapping task with step-change perturbations. The experiment was conducted at the Department of Science and Technology, University of Quilmes in Bernal, Buenos Aires, Argentina. The participants were 10 volunteers (one female, ages 18-36), with no training in finger-tapping tasks. All participants had substantial musical training (7 years or more; only two had less than 3 years). Five of them played percussion.

### 3.2. Materials and equipment

The recording of finger tapping with a standard keyboard and computer has a number of drawbacks, such as delays due to multitasking operating systems, delays due to the keyboard buffer, the key elastic response to tapping, full key depression versus key oscillation, etc. To overcome these difficulties we designed and built an electronic transducer which communicates both ways with a desktop computer through a stereo full-duplex sound card sampling at 8 kHz, yielding a time resolution of 0.125 ms (for either the presentation of the stimuli or the recording of the responses). Participants tapped using a small electrode attached to their finger's tip on a rigid ground copper plate (the subjects were electrically isolated). The small electrode is connected through a resistor to a 9V battery, so that the electronic transducer detects the tap by the change in electrode voltage. The voltage drop triggers an LM555 timer operating in monostable mode which produces a standardized square pulse of 50 ms duration; the pulse is then registered as a sound signal in one channel of the sound card. Sound feedback from the tap is produced by a second LM555 in astable mode (600 Hz) using the output of the first timer as the envelope. The stimulus is represented by a similar logic signal that is played as a sound file through the sound card. The transducer then transforms the logic signal into an audible bip by operating a similar tandem of LM555 timers in monostable (sound envelope) and astable (sound oscillation) modes, generating pulses of 320 Hz frequency and 50 ms total duration. The envelope of the stimulus signal is registered simultaneously on the second channel of the sound card so that any delay either in the recording software or the operating system or the electronic transducer affects equally both channels, leaving the asynchronies unaltered. The delays produced by the electronic processing in the transducer were measured and found to be smaller than 50  $\mu$ s. Recording and playback of logic signals through the sound card were performed using Audacity 1.2.6, which is a free, open-source software for recording and editing sounds. The data were analyzed using custom code in MATLAB. Printed-circuit board and detailed list of components is available from the authors on request.

We chose to use auditory feedback from the taps in order to have a reduced timing variability (Drewing and Aschersleben, 2003), which was important to show the subtle asymmetry effect in the smaller perturbation magnitudes; and to have a more realistic musical context, as one of our goals was to include the perturbation within the modeling effort which is relevant for musical contexts and interpersonal synchronization. This is a difference between our work and some of the works we cited, e.g. Repp (2001b, 2002a).

### 3.3. Procedure and data

Participants listened over headphones at a comfortable intensity and tapped with their index finger on the copper plate. Each trial consisted of a sequence of 60 tones (pre-perturbation interstimulus interval 500 ms) and was perturbed only once, randomly between the 20th and the 30th tone. There were 5 perturbation magnitudes (10, 20, 30, 40 and 50 ms) and 2 perturbation directions (positive and negative).

Each tempo change was presented 5 times. Each participant was presented with 50 sequences (5 magnitudes x 2 directions x 5 repetitions) in random order. Participants were instructed to synchronize their taps with tones as accurately as possible. They were alerted that the tone sequence might change tempo at a random position. After each trial, they were asked to report if they were aware of any perturbation.

We recorded the absolute times of occurrence of every tone ( $S_n$ ) and tap ( $R_n$ ) as described above, and took the differences  $e_n = R_n - S_n$ . For each condition (perturbation magnitude  $\times$  direction) we aligned the data at  $n = 0$  and average within subject, then averaged across subjects. All experimental data shown in this work represent mean across subjects  $\pm$  standard error of the mean across subjects. To exclude irrelevant variation due to interindividual differences in the pre-perturbation baseline, we computed the pre-perturbation baseline as the average of the ten asynchronies before the perturbation ( $-9 \leq n \leq 0$ ) for each condition and then subtracted this value from all asynchronies in the corresponding time series, such that the pre-perturbation baseline is zero.

### 3.4. Post-perturbation baseline

We observed a systematic shift of the baseline (the point of subjective synchrony) before-after the perturbation. This is a novel result discussed below, and was quantified by taking the average of the last 10 asynchronies of the corresponding experimental time series for each condition and subject (after relativizing to the corresponding pre-perturbation baseline as described above), then averaged across subjects.

### 3.5. Genetic algorithm and model simulations

The fitting of the model was performed by a genetic algorithm in C, using both custom-written code and the GAUL libraries (<http://gaul.sourceforge.net>). Numerical simulations were performed with custom code in C. Code files are available from the authors on request.

The choice of a genetic algorithm was based on the following reasoning. It is known that the system to be optimized, if nonlinear, could undergo a bifurcation after the fitting algorithm changes the value of any parameter, what could be reflected in a discontinuity in the fitness function. This is an important shortcoming for gradient-based algorithms, as all gradient algorithms are ill-defined for non-differentiable objective functions. Additionally, gradient descent and methods like Newton or Gauss-Newton by their own definitions can only find the closest local optimum, a particularly important shortcoming when dealing with a nonlinear model in a 7-dimensional parameter space like ours.

On the other hand, genetic algorithms perform an efficient search across the entire parameter space first (as defined by the initial parameter ranges), find a good global solution, and then find the local optimum around it. While it is true that sometimes a genetic algorithm can converge to a local optimum (as many methods do, other than brute force exhaustive search), this can be avoided by several “countermeasures” like for instance just increasing the mutation rate, or performing the evolution several times and then choosing the absolute optimum (the chosen strategy in our case).

The seven model parameters were arranged into one chromosome with seven genes, and initialized randomly with a uniform distribution in the ranges  $-1.0 < a, b, c, d < 1.0$  and  $-0.01 < \alpha, \beta, \gamma < 0.01$ , which almost exhausts the possible combinations of meaningful values (larger values for the linear coefficients for instance would mean that the variable  $e_n$  be corrected by a proportion greater than 100% of the previous value; also, large values would most likely yield eigenvalues greater than 1, with the unrealistic result that the time series would be repelled away from the point of subjective synchrony).

The fitness function was defined as minus the square root of the average squared deviation between model series and experimental series at each step. That is, if the experimental time series is  $E_n^j$  and the model time series is  $e_n^j$  ( $n = 1, \dots, 20$  is the step number,  $j = 1, 2$  represents the two conditions  $\pm 50$  ms used to fit the model), then the fitness function  $F$  reads

$$F = -\sqrt{\frac{1}{40} \sum_{j=1}^2 \sum_{n=1}^{20} w_n^j (E_n^j - e_n^j)^2 + P}$$

which decreases as the differences  $E_n^j - e_n^j$  get larger in absolute value. The weights  $w_n^j$  were chosen to counteract the effect of the shallower, longer approach to equilibrium and give more importance to the quicker, shorter transient part right after the perturbation:

	$w_n^{j=1}$	$w_n^{j=2}$
$1 \leq n \leq 3$	15	10
$4 \leq n \leq 8$	30	10
$9 \leq n \leq 20$	1	1

In order to prevent survival of unrealistic solutions (for instance damped oscillations, as discussed in section 2.3), penalties were included as a positive term  $P$  inside the square root that depends on the linear coefficients only and takes a large value in any of the following cases:

1. the eigenvalues are complex (in order to avoid oscillatory approach to the equilibrium);
2. the eigenvalues are real but any of them is either greater than 1 or negative (in order to avoid solutions that diverge from the equilibrium, and convergent solutions that alternate sides);
3. the eigendirections have slopes with same sign (in order to enforce a generic approach to the equilibrium from each side).

The algorithm was stopped at 50 generations, which allowed the fitness function to reach a constant value. Crossover rate and mutation rate were set at the usual values of 0.9 and 0.1, respectively, and the population size was 500. In order to prevent the selection of a surviving local optimum, the whole procedure described so far was repeated 100 times; the chosen solution was the one with the highest fitness of all ( $F = -12.35$ ). Subsequent analysis of the set of 100 partial solutions revealed a unimodal distribution for every parameter estimate, building confidence on the global nature of the chosen solution. The converged

linear coefficients are of the same order as previously published values (see for instance Repp (2001a); Thaut et al. (1998); Schulze et al. (2005)).

The model was fitted to the time series data from our step-change perturbation experiment, using the  $\Delta = \pm 50$  ms time series only (i.e. only the extreme perturbations). To improve fitting, a constant baseline was added to the model variable  $e_n$  after the simulated perturbation, with a fixed value equal to the post-perturbation baseline. For reasons discussed below, we fitted our model to the experimental time series for steps  $n \geq 1$ .

### 3.6. Amount of overshoot $\delta$

The amount of overshoot  $\delta$ , which is a function of the perturbation size, was defined for the positive perturbations as the difference between the maximum of the time series and the post-perturbation baseline for each condition; the error is the error of the maximum (zero in the case of the model). For negative perturbations, the amount of overshoot  $\delta$  was defined likewise but using the minimum of the time series instead. We used the same definition for both experimental data and model simulations. Note that this definition, although very straightforward, is also very crude and can yield a value of  $\delta$  different from zero even if there is no overshoot, provided that the time series are noisy. To see this, consider a finger tapping experiment to an isochronous metronome with no perturbations—since the time series is naturally noisy, the difference between the maximum (or the minimum) of the time series and the average value will obviously always be different from zero despite being a stationary time series (i.e. no overshoot at all). This bias is particularly detrimental for our negative perturbations, where  $\delta$  overestimates (in absolute value) an overshoot that can hardly be seen. The behavior we report, however, is robust enough to display a clear quantitative asymmetry despite this bias and so we stick with this straightforward definition.

## 4. Experimental Results

### 4.1. Step-change perturbations

Our experimental time series are displayed in Figure 6, along with simulations of our model using the parameter values fitted in Figure 5B (see Table 1 and Materials and Methods).

The qualitative agreement between data and model is remarkable, particularly because all simulations were performed with a single set of parameter values and because the model was fitted to the  $\pm 50$  ms perturbations only. The response to positive perturbations (top row) displays an overshoot which is larger for larger perturbation magnitudes, while for the negative perturbations (bottom row) the overshoot is almost undetectable, both features as predicted by our model.

There are quantitative differences worth noticing, though, which can be seen in the right column of Figure 6. While the overshoot is well reproduced at the two largest perturbations ( $\Delta = +50$  and  $+40$  ms, compare panels A and B), in the experimental data it appears to develop more slowly for smaller perturbations which is evidenced in panel C by a larger difference at steps  $n = 0$  through  $n = 5$  (see the awareness data below).

The main prediction of our model is shown in Figure 7, which displays the amount of overshoot  $\delta$  as a function of the perturbation size  $\Delta$  (see Materials and Methods). Despite the quantitative differences described above between the time series, the correspondence between the experimental data and model simulations is remarkable (recall that the only time series that were used to fit the model correspond to the extreme points in this plot,  $\Delta = +50$  and  $-50$  ms). Both the experiment and the model display a very similar nonlinear relationship between overshoot  $\delta$  and perturbation size  $\Delta$ , with a flat profile at negative perturbations and a smooth, accelerated increase at positive perturbations. We want to point out that the overshoots for the negative perturbations, which in the model are almost zero, might even be overestimated for the experimental data, and might be actually closer to zero; see Materials and Methods for details on the definition of  $\delta$ .

Note the shape of the response immediately following the perturbation ( $n = 1$ ). It takes one step—or perhaps two steps at the smallest magnitudes—for the subjects to fully develop their response to a step change, both for positive and negative perturbations. This behavior could be already noticed in the work of Thaut et al. (1998); see Figure 3B in this work, at  $n = 1$ . Our model does not account for this feature—or any other published model to the best of our knowledge unless the parameter values are changed for each condition (tempo, perturbation size, etc.). This phenomenon has been considered evidence suggesting two different correction strategies operating simultaneously—namely period and phase correction, first proposed by the two-process model of Mates (1994a,b). More parsimoniously, we interpret this as either the subject being aware that a perturbation happened but not immediately interpreting it as a change in stimulus period, or not being able to correct completely despite being aware of the perturbation. More generally, the hypothesis that the neural mechanisms underlying the correction behavior do not fully develop until at least the second tap cannot be ruled out at this point. Accordingly, we fitted our model to the experimental time series for  $n \geq 1$ . A more complete description of this behavior is under development.

#### *4.2. Phase-shift and event-onset shift perturbations*

Our model is also capable of reproducing the observed behavior following phase-shifts and event-onset shifts. This should be considered as a prediction too, since we didn’t take into account any of these two perturbations in the building of the model, except for the observation that subliminal phase shifts do not display threshold (Section 2.2.1). In the same line of reasoning as above for the reproduction of the step changes, we treat phase-shift and event-onset shift perturbations as follows.

In a phase-shift perturbation, the interstimulus interval before and after the perturbation has the same value, where the main effect is the resetting of the variable  $e_0 \rightarrow e_0 - \Delta$  due to the unexpected “shift” of the stimulus at  $n = 0$  (Section 2.3.1). We base our modeling on the observation that the response of the subjects to a step change takes at least one step to fully develop (Section 4.1), together with the fact that in a phase shift the interstimulus interval returns to the original value after two steps (Section 2.3.1). We

then assume that the only effect of this perturbation on our model is the resetting of the variable  $e_n$  at step  $n = 0$  with no change in  $T_n$ . The subsequent evolution of  $e_n$  for  $n \geq 1$  is given by the model's equations.

In an event-onset shift perturbation, the interstimulus interval returns to the original value after three steps. Consistently, we assume that the only effect on our model is the resetting of the variable in two steps, with no change in  $T_n$ . At the first step of the perturbation,  $n = 0$ , the variable is changed  $e_0 \rightarrow e_0 - \Delta$ . At the second step  $n = 1$  the variable should be reset again in the opposite direction  $e_1 \rightarrow e_1 + \Delta$ . Notice that this does not mean that the consecutive perturbations should cancel each other, because the correction mechanism is already engaged at step  $n = 1$  and thus  $e_1 \neq e_0 - \Delta$ . Subsequent evolution of  $e_n$  for  $n \geq 1$  is always driven by the model's equations.

The correspondence between the experimental traces and our model's traces after both phase shifts and event-onset shifts is remarkable, as can be seen in Figure 8. The only easily noticeable difference between model simulation and the corresponding experimental series is in the +50-ms event-onset shift (panel D), although they have the same qualitative behavior. Note that the fitting of the model did not make use of any perturbation other than a step change.

#### 4.3. Individual data

Figure 9 displays the experimental time series of four subjects, averaged across repetitions. There are some noticeable interindividual differences; for instance, subject MA displays the slowest return to the baseline, both for positive and negative perturbations. Subject PA displays the smallest range of post-perturbation baselines for positive perturbations, but the largest range for negative perturbations. The overall qualitative behavior, however, is common to all subjects.

The right column in Figure 9 displays the amount of overshoot  $\delta$  as a function of perturbation size  $\Delta$  for the individual subjects. Note the same asymmetry for all subjects—positive perturbations display larger overshoot than negative perturbations. There are a few points at negative values of  $\Delta$  that are noticeably below zero (e.g. subject RU,  $\Delta = -50$  ms). Recall, however, that our definition of  $\delta$ , although straightforward, is biased and can actually yield a value different from zero in the absence of overshoot if the time series is naturally noisy (see Materials and Methods).

#### 4.4. Shift in post-perturbation baseline and effect of awareness

Another novel finding is a systematic shift of the baseline before-after the perturbation, shown in Figure 10A. Note first the asymmetric dependence on the perturbation size: positive perturbations lead to smaller (in absolute value) post-perturbation baselines than negative perturbations. Note also that the baseline in the  $\pm 20$  and  $\pm 10$  ms perturbations is very close to zero. The shift in pre-post baseline could also be seen in the step-change data reported by Repp (2001b, Figure 4), although in his work even the smallest perturbation sizes seem to display the shift. This difference is probably due to the fact that we used auditory

feedback from the taps; if this hypothesis holds, then the behavior of the baseline shift for larger perturbation sizes is a true effect of the perturbation.

Figure 10B displays the fraction of correct responses (detection of whether a perturbation occurred or not) as a function of the perturbation size. Note that the smallest perturbations ( $\pm 10$  ms) were just below the detection threshold (0.5) and thus were undetectable most of the times. Even the  $\pm 20$  ms perturbations were near the detection threshold, though above it. This could explain the slower development of the overshoot after the positive perturbations (Figure 6) at the smallest perturbation magnitudes.

## 5. Discussion and Conclusions

Although pioneering work on timing goes back as far as the mid 20th century, e.g. Michon (1967), the representation of time in our brain has only begun to rise as a fundamental issue in neuroscience (Ivry and Schlerf, 2008). During the last decade, several studies have been published on the theoretical and experimental aspects of neural timing (Ivry and Schlerf, 2008; Mauk and Buonomano, 2004; Buhusi and Meck, 2005) and sensorimotor synchronization (for a thorough review see Repp (2005)). Still, our understanding of how the brain discriminates between two interval durations or how the brain produces precisely timed motor responses is far from clear.

In the case of sensorimotor synchronization, where several, probably distinct neural systems are recruited—most notably sensory and motor timing—, a phenomenological behavioral approach has proven to be fruitful in the study of the phenomenon (Repp, 2005). This common approach groups into a single model the representation of several different neural processes, namely the perception of time intervals, comparison and decision making, and the production of timed motor responses. Detailed theoretical modeling of the neural bases for time estimation in the range of hundreds of milliseconds is an active area of research, both for sensory and motor timing (for reviews see Buonomano and Laje (2010) and Ivry and Schlerf (2008)); many central issues such as whether sensory and motor timing rely on the same circuitry still remain open. Thus, the particular features that characterize the sensorimotor synchronization behavior—distinguishing it from either pure time perception tasks, or pure decision making tasks, or pure time production tasks—make the phenomenological behavioral approach a very valuable one, at least until we reach a deeper understanding of its internal workings.

In our work we thus chose a phenomenological behavioral approach, in the spirit of searching for the dynamical constraints that the observed behavior sets on the underlying neural system. Any neural model proposed to account for the behavior should take those constraints into consideration. As an analogy, consider a system where the experimenter has only access to a stroboscopic measure of some observable, rather than to the actual continuous underlying variables—any detailed model of the underlying system should have a Poincaré section that resembles the dynamics of the stroboscopic measure.

### 5.1. Nonlinear behavior

We showed that the behavior consistently displays nonlinear effects in the form of an asymmetric response following a small step-change perturbation (smaller than 10% of the period), whereas common nonlinear effects like saturation appear at larger perturbation magnitudes (25-50%) (Repp, 2002b). Figure 7 provides strong evidence of nonlinear effects at small perturbation magnitudes, where the behavior is traditionally represented by linear models.

We do not deny that the overshoot may occur after negative step changes if larger perturbation magnitudes are considered. However, it is clear from Figure 7 that, for perturbation magnitudes up to  $\pm 50$  ms, the overshoot only occurs after positive perturbations. This is sufficient evidence for claiming an asymmetric behavior in a self-consistent way within a small-perturbation, second-order approach like ours. We purposely decided to restrict our study to small perturbations, and accordingly include only terms up to second order, in order to keep the nonlinearity in the data to a minimum that could be accounted for in a consistent way. Even if an overshoot were eventually found after large negative perturbations, any future model would still have to reproduce the asymmetric features shown here for smaller perturbation magnitudes.

It could be argued that the asymmetry is related to the computational costs for correcting positive and negative asynchronies. The costs might differ, because correcting a step decrease in stimulus period would necessitate shortening the interresponse interval accordingly and thus carry a loss of processing time (Aschersleben, 2002). However, this is not consistent with the experimental data shown in Figure 6: either with the apparently straightforward recovery after a step decrease—a monotonic exponential convergence to the baseline—, or with the overshoot only following a step increase—where the subject would have indeed more time for computational purposes.

Repp and Keller (2004) demonstrated that the fits of the dual-process Mates' model (Mates, 1994a,b) to data with tempo perturbations changed significantly contingent on the subject's awareness. Although our model offers qualitatively successful description and predictions of the behavior, the lack of a quantitative fit right after the perturbation for the smallest perturbation sizes (Figure 6C and F) points to the likely need to address awareness issues. They also stated, "The crucial assumption of the two-process model is that phase correction and period correction are independent of each other, so that their behavioural effects are additive", and "phase correction and period correction seem to represent independent processes". The need of quadratic terms demonstrated in our work, however, speaks against the strictly additive effects, and calls for a revision of the idea of independent processes (particularly because of the cross-term  $\alpha e_n x_n$ , see below).

We proposed a simple (quadratic) model that reproduces a great variety of qualitative behaviors in a finger-tapping task. From a dynamical system's perspective, there is a clear-cut ranking for simplicity: what is the order of the highest-order term in the model? In this sense, a quadratic model is the next-to-simplest model. On the other hand, a function like tanh for instance (usually representing saturation) can be very

compact but in fact it can be considered as a placeholder for an infinite sum of terms with any desired odd power ( $\tanh(x) \sim x + ax^3 + bx^5 + cx^7 + \dots$ ). Our model is indeed a qualitative, not quantitative, model of the dynamics. In view of the step-change data, however, this is much stronger than just adding an additional term to improve the fitting by certain amount. Why would “qualitative” be stronger than “quantitative”? Because we point out to a *qualitative difference* between the data and any linear model—not just a quantitative one. The qualitative difference refers to the fact that the system displays an asymmetric response (overshoot for positive perturbations only), which cannot be described by any linear model. It is simply not possible for a linear model to respond asymmetrically to a symmetric perturbation—it will show an overshoot either to both perturbations or to none (Loehr et al., 2011). Of course any good model should eventually provide a quantitative fit, and this is our first step towards it.

The interpretation of the nonlinear terms should be approached with care. Within the traditional dual-process framework (Mates, 1994a,b), the internal timekeeper is updated by period and phase correction and thus the cross term  $\alpha e_n x_n$  would be interpreted as an interaction between the two processes. This interaction could be due for instance because of shared neural substrates or cross-input between different substrates (see for instance Praamstra et al. (2003)). This does not mean that an interaction between the period and phase correction processes is necessary for the internal timekeeper to work well or to be precise. Actually, the experimentally observed asymmetry and overshoot can hardly be considered as “desirable” features of any proposed internal timekeeper—perhaps this is just the way it is.

Within a more general framework, where the variable  $x_n$  might be interpreted as either a measure of the subject’s perception of the interstimulus interval duration (or the subject’s estimate of the next interstimulus interval, or even the subject’s prediction of his/her own next interresponse interval—see below), and the variable  $e_n$  as a measure of the subject’s perception of the asynchrony, the cross term  $\alpha e_n x_n$  could be interpreted as an interaction between the two percepts, or also as follows. Recall eq. (5) and note that the terms  $e_{n+1} = ae_n + \alpha e_n x_n + \dots$  can be regrouped as  $e_{n+1} = (a + \alpha x_n)e_n + \dots$ . This helps interpreting the nonlinear cross term as a correction to the linear term; that is, the coefficient of  $e_n$  is not constant but varies, and its value is modified by the measure  $x_n$  of the subject’s internal perception of the interstimulus interval. In much the same way, the term  $\beta e_n^2$  can be interpreted as a correction to the coefficient of the linear term  $ae_n$ —thus the amount of correction coming from the perception of an asynchrony would be disproportionally larger for larger asynchronies.

Further work is needed to unveil the true nature of the nonlinear terms. We found, for instance, some parameter interdependence: the estimates for  $\alpha$  and  $\beta$  are negatively correlated (data not shown). This means that two different interpretations can account for the data in much the same amount: either with a larger  $\beta$  and a smaller  $\alpha$  (and thus with a more important interaction between the two variables), or the other way around (and thus a more important “auto-interaction” for  $e_n$ ). Quadratic terms, however, are necessary to reproduce the observed behavior anyways.

The systematic shift of the post-perturbation baseline with respect to its pre-perturbation value was an unexpected finding (Figure 10A), also pointing to the asymmetric features of the behavior but unexplained by our model. It is known that the NMA for auditory sequences depends on the interstimulus interval (Repp, 2003), displaying a shallow decrease spanning less than 20 ms as a function of the interstimulus interval in the range 320–680 ms (i.e. it becomes more negative for slower sequences). The NMA after the perturbation could change because the interstimulus interval changes. However, our result is not consistent with that: the post-perturbation baseline in our experiment actually increases as a function of the perturbation size (i.e. it becomes more positive for slower sequences). It seems to be dependent on awareness (because it only appears after perturbations larger than  $\pm 20$  ms), but it seems to be also dependent on the perturbation size (i.e. not constant) if the perturbation is above the detection threshold. Future work should either correct this finding or reconcile it with the accepted causes and explanations for the NMA (e.g. a slower central registration of tactile and proprioceptive information as compared with the auditory modality, or the phase lag in coupled oscillator models; see Repp (2005) for a review). The fact that we used auditory feedback from the taps is one likely cause for the absence of NMA at the smallest perturbations ( $\pm 20$  and  $\pm 10$  ms); the remaining points seem to be a pure effect of the perturbation.

### 5.2. *Single vs. several strategies*

Beyond the mere fitting of a new model, we proposed a unified view for the effect of a perturbation on the behavior. This unified framework is a first step to consistently include the perturbation within the modeling effort (Thaut et al., 1998), an important goal when interpersonal or crowd synchronization is studied, as in music, or in computer-human synchronization experiments. Moreover, our model takes into account the three most common finger-tapping perturbations on equal footing. Previous studies have fitted a different set of parameter values to each perturbation magnitude (see e.g. Thaut et al. (1998); Repp (2001b); Schulze et al. (2005)). We emphasize that our model correctly predicts the relationship between overshoot and perturbation size for the step-change perturbations, and qualitatively predicts the evolution of the observed asynchronies following all studied perturbation types and magnitudes with a single set of parameter values.

The sensorimotor synchronization behavior is likely to draw on several distinct neural processes, namely time perception, interval comparison, error detection, time production, and motor execution. However, the question remains open as to whether this superposition leads to different strategies for different perturbations, or whether the whole behavior can be interpreted as the result of a single timing mechanism/strategy. The simplicity of our proposed unified framework and the success of our model point to the singleness of the underlying mechanism responsible for the synchronization.

Assume for a moment that the parameter values are to be changed for each perturbation magnitude or direction. As we discussed above, this would mean considering different strategies, each represented by a different parameter value set; for instance in the period vs. phase correction framework, period correction

could have a greater relative importance for some perturbation magnitudes whereas phase correction could for some others. Additionally, a complete description would need a supplementary timing mechanism to quickly choose among all the parameter sets available (strategies) the one that applies to the current perturbation, as the subject only knows the perturbation magnitude or nature once it is gone. We emphasize that the parameter values in our work are fixed and thus there is no possibility of switching between strategies, and that our model reproduced the behavior after the three perturbation types without resorting to any additional “strategy switching” mechanism.

The different terms in our model might or might not correspond to different underlying neural processes. A key issue in our proposal, however, is that we do not turn them on and off depending on the perturbation magnitude or sign, or change the value of their coefficients. The two equations that constitute our model are always “on” and with constant parameter values. It is in this sense that our model points to a single mechanism—there are no “different strategies” given by different sets of parameter values; there is no need to resort to additional mechanisms whose task would be to select the appropriate value of the parameters; and there is a fixed set of terms in the model.

We also showed that the behavior following a step-change perturbation cannot be explained in terms of a single variable only, namely the observable  $e_n$ , due to the observed overshoot. The overshoot is a pretty robust behavior after step increases, and can be noticed also in Figure 4B of Repp (2001b), after removing the pronounced linear drift displayed by the subjects after the perturbation. We then argued for the need of a second variable,  $x_n$ , in principle of purely dynamical origin, whose evolution is given by an equation that resembles the so-called “period correction” mechanism (Repp, 2005). Several previous studies have proposed two distinct correction processes, namely phase and period correction mechanisms, based on the hypothesis that there is an internal timekeeper or oscillator whose parameters can be reset or modified (see for instance Mates’ influential timekeeper model (Mates, 1994a,b), and Large’s oscillator model (Large, 2000)). We avoid such interpretations and, perhaps more parsimoniously, propose that our variable  $x_n$  might be interpreted as either a measure of the subject’s perception of the interstimulus interval duration, or the subject’s estimate of the next interstimulus interval, or even the subject’s prediction of his/her own next tap (based on either the alleged internal oscillator or a more complex neural construct). A more complex, although probably true, interpretation is that  $x_n$  is a compound measure of all. It is difficult to base interpretations on the relationship between model parameters and the underlying neural/psychological system without further experimental manipulation, particularly because of the composite nature of this behavior. Further investigation will be necessary to establish the neural/psychological identity of the dynamical variable  $x_n$ .

Phase shifts and event-onset shifts are usually regarded as more primitive and simple forms of perturbation than step changes (Repp, 2005). This could be due to the smaller perceptual impact on the subject; indeed, if no correction were made, the asynchrony after a phase shift would remain constant, and after an event-onset shift there would be no asynchrony at all, whereas the asynchrony after a step change would

increase linearly with time. However we argue for a different interpretation. Since in our framework the step change perturbation involves only one change in parameter value ( $T \rightarrow T + \Delta$ ), while the phase shift involves two changes ( $T \rightarrow T + \Delta \rightarrow T$ ) and the event-onset shift involves three changes ( $T \rightarrow T + \Delta \rightarrow T - \Delta \rightarrow T$ ), we argue in favor of considering the step change a more fundamental perturbation. Our model consistently incorporates this interpretation. The remarkable result, however, is not that the model can be driven through phase shifts or event-onset shifts as a concatenation of step changes; what we find remarkable is that the experimental data from the other perturbation types are correctly predicted as a concatenation of step changes (Figure 8). Up to our knowledge, it is not known whether the neural mechanisms/strategies underlying the response to a step change are the same as those underlying a phase shift or an event-onset shift. However, we succeeded in predicting both the experimental phase shift and event-onset shift as a concatenation of responses to individual step changes, which in addition suggests that the underlying system might be considered as a single mechanism/strategy.

Our model was not designed to reproduce the small drift observed immediately after the smallest step-change perturbations: it takes 2 or 3 steps to revert the trend and make the asynchrony actually decrease (Figure 6A and C). Error correction, however, is engaged already at the first step immediately after the perturbation—otherwise the time series would follow a straight line with slope  $-\Delta$ . The possibility of at least some subjects having figured out that all perturbations were step changes cannot be ruled out. Future work should address this particular behavior, and also the response to larger perturbation magnitudes and other base tempos.

## 6. Acknowledgements

We thank Dean V. Buonomano, Mariano Sigman, Juan Kamienkowski, Manuel C. Eguía, Andrés Benavides, and Guillermo Solovey for helpful discussions and comments on the manuscript. This work was supported by CONICET and ANPCyT PICT-2007-881 (Argentina), and the Fulbright Commission (USA-Argentina).

## References

- Aschersleben, G., 2002. Temporal control of movements in sensorimotor synchronization. *Brain and Cognition* 48, 66–79.
- Beudel, M., Renken, R., Leenders, K. L., de Jong, B. M., 2009. Cerebral representations of space and time. *NeuroImage* 44, 1032–1040.
- Buhusi, C. V., Meck, W. H., 2005. What makes us tick? functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience* 6, 755–765.
- Buonomano, D. V., Laje, R., 2010. Population clocks: motor timing with neural dynamics. *Trends in Cognitive Sciences* 14, 520–527.
- Chen, Y., Ding, M., Kelso, J. A. S., 1997. Long memory processes ( $1/f^a$  type) in human coordination. *Physical Review Letters* 79, 4501–4504.
- Del Olmo, M. F., Cheeran, B., Koch, G., Rothwell, J. C., 2008. Role of the cerebellum in externally paced rhythmic finger movements. *Journal of Neurophysiology* 98, 145–152.
- Drake, C., Botte, M.-C., 1993. Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception and Psychophysics* 54, 277–286.
- Drewing, K., Aschersleben, G., 2003. Reduced timing variability during bimanual coupling: a role for sensory information. *The Quarterly Journal of Experimental Psychology* 56A, 329–350.

- Engbert, R., Krampe, R. T., Kurths, J., Kliegl, R., 2002. Synchronizing movements with the metronome: Nonlinear error correction and unstable periodic orbits. *Brain and Cognition* 48, 107–116.
- Gibbs, C. B., 1965. Probability learning in step-input tracking. *British Journal of Psychology* 56, 233–242.
- Gross, J., Timmermann, L., Kujala, J., Dirks, M., Schmitz, F., Salmelin, R., Schnitzler, A., 2002. The neural basis of intermittent motor control in humans. *Proceedings of the National Academy of Science USA* 99, 2299–2302.
- Hary, D., Moore, G. P., 1987a. Synchronizing human movement with an external clock source. *Biological Cybernetics* 56, 305–311.
- Hary, D., Moore, G. P., 1987b. Temporal tracking and synchronization strategies. *Human Neurobiology* 4, 73–77.
- Hirsh, I. J., 1959. Auditory perception of temporal order. *Journal of the Acoustical Society of America* 31, 759–767.
- Hirsh, I. J., Scherrick, C. E., 1961. Perceived order in different sense modalities. *Journal of Experimental Psychology* 62, 423–432.
- Ivry, R. B., Schlerf, J. E., 2008. Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences* 12, 273–280.
- Ivry, R. B., Spencer, R. M. C., 2004. The neural representation of time. *Current Opinion in Neurobiology* 14, 225–232.
- Large, E. W., 2000. On synchronizing movements to music. *Human Movement Science* 19, 527–566.
- Large, E. W., Fink, P., Kelso, J. A. S., 2002. Tracking simple and complex sequences. *Psychological Research* 66, 3–17.
- Large, E. W., Jones, M. R., 1999. The dynamics of attending: how people track time-varying events. *Psychological Review* 106, 119–159.
- Lewis, P. A., Miall, R. C., 2003. Brain activation patterns during measurements of sub- and supra-second intervals. *Neuropsychologia* 41, 1583–1592.
- Loehr, J. D., Large, E. W., Palmer, C., 2011. Temporal coordination and adaptation to rate change in music performance. *Journal of Experimental Psychology: Human Perception and Performance* 37, 1292–1309.
- Luce, R. D., 1986. Response times. Oxford University Press.
- Madison, G., Merker, B., 2004. Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neuroscience Letters* 370, 69–73.
- Manto, M., Bastian, A. J., 2007. Cerebellum and the deciphering of motor coding. *The Cerebellum* 6, 3–6.
- Mates, J., 1994a. A model of synchronization of motor acts to a stimulus sequence: I. timing and error corrections. *Biological Cybernetics* 70, 463–473.
- Mates, J., 1994b. A model of synchronization of motor acts to a stimulus sequence: II. stability analysis, error estimation and simulations. *Biological Cybernetics* 70, 475–484.
- Mauk, M. D., Buonomano, D. V., 2004. The neural basis of temporal processing. *Annual Review of Neuroscience* 27, 307–340.
- McAuley, J. D., Kidd, G. R., 1998. Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance* 24, 1786–1800.
- Meck, W. H., 2005. Neuropsychology of timing and time perception. *Brain and Cognition* 58, 1–8.
- Michon, J. A., 1967. Timing in temporal tracking. Assen, The Netherlands: van Gorcum.
- Panda, S., Hogenesch, J. B., Kay, S. A., 2002. Circadian rhythms from flies to human. *Nature* 417, 329–335.
- Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., Perryer, L., 2003. Neurophysiological correlates of error correction in sensorimotor-synchronization. *NeuroImage* 20, 1283–1297.
- Pressing, J., 1998. Error correction processes in temporal pattern production. *Journal of Mathematical Psychology* 42, 63–101.
- Pressing, J., Jolley-Rogers, G., 1997. Spectral properties of human cognition and skill. *Biological Cybernetics* 76, 339–347.
- Repp, B., 2000. Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychological Research* 63, 106–128.
- Repp, B., 2001a. Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance* 27, 600–621.
- Repp, B., 2001b. Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Human Movement Science* 20, 277–312.
- Repp, B., 2002a. Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance* 28, 410–430.
- Repp, B., 2002b. Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science* 21, 1–37.
- Repp, B., 2005. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review* 12, 969–992.
- Repp, B. H., 2003. 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* 4, 355–370.
- Repp, B. H., 2008. Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research* 72, 79–98.
- Repp, B. H., 2011. Tapping in synchrony with a perturbed metronome: The phase correction response to small and large phase shifts as a function of tempo. *Journal of Motor Behavior* 43, 213–227.
- Repp, B. H., Keller, P. E., 2004. Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology* 57A, 499–521.
- Schachner, A., Brady, T. F., Pepperberg, I. M., Hauser, M. D., 2009. Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology* 19, 831–836.
- Schöner, G., 2002. Timing, clocks, and dynamical systems. *Brain and Cognition* 48, 31–51.
- Schulze, H.-H., 1992. The error correction model for the tracking of a random metronome: statistical properties and an empirical test. In: Macar, F., Pouthas, V., Friedman, W. J. (Eds.), *Time, action, and cognition: Towards bridging the gap*. Dordrecht: Kluwer, pp. 275–286.
- Schulze, H.-H., Cordes, A., Vorberg, D., 2005. Keeping synchrony while tempo changes: accelerando and ritardando. *Music*

- Perception 22, 461–477.
- Semjen, A., Schulze, H.-H., Vorberg, D., 2000. Timing precision in continuation and synchronization tapping. *Psychological Research* 63, 137–147.
- Semjen, A., Vorberg, D., Schulze, H.-H., 1998. Getting synchronized with the metronome: comparisons between phase and period correction. *Psychological Research* 61, 44–55.
- Thaut, M. H., Miller, R. A., Schauer, L. M., 1998. Multiple synchronization strategies in rhythmic sensorimotor tasks: phase vs period correction. *Biological Cybernetics* 79, 241–250.
- Vorberg, D., Schulze, H.-H., 2002. Linear phase-correction in synchronization: predictions, parameter estimation, and simulations. *Journal of Mathematical Psychology* 46, 56–87.
- Wagenmakers, E.-J., Farrell, S., Ratcliff, R., 2004. Estimation and interpretation of  $1/f^\alpha$  noise in human cognition. *Psychonomic Bulletin & Review* 11, 579–615.
- Wing, A. M., Kristofferson, A. B., 1973. Response delays and the timing of discrete motor responses. *Perception and Psychophysics* 14, 5–12.

### Figure and table captions

$a = 0.408$	$\alpha = -0.0065 \text{ ms}^{-1}$
$b = -0.099$	$\beta = -0.0016 \text{ ms}^{-1}$
$c = -0.052$	$\gamma = 0.015 \text{ ms}^{-1}$
$d = -0.266$	$T = 500 \text{ ms}$

Table 1: Model parameter values used throughout this work. Linear coefficients  $a, b, c, d$  are nondimensional.  $T$  is the pre-perturbation interstimulus interval.

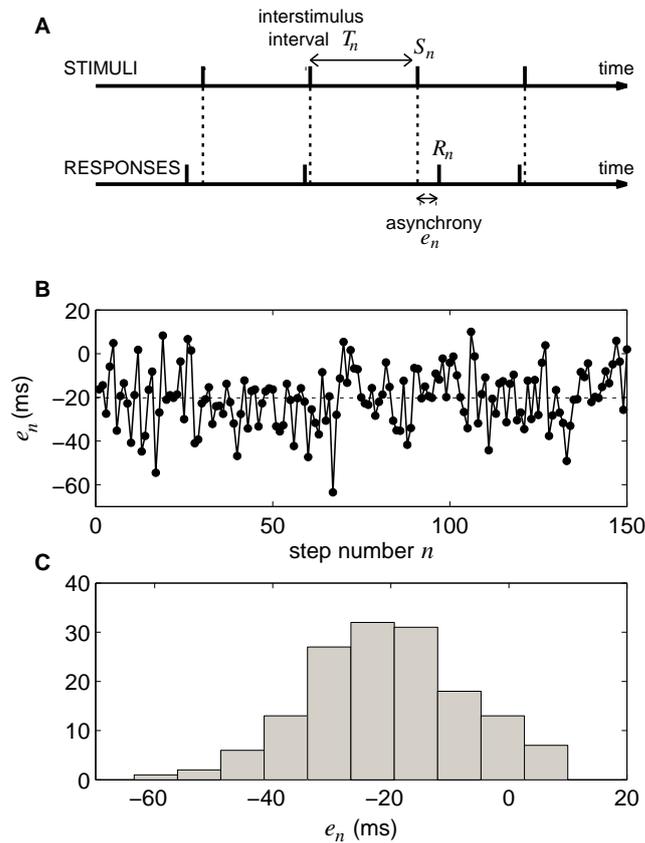


Figure 1: The finger-tapping task. (A) Definition of variables. The subject is instructed to tap in synchrony with a metronome, and the differences  $e_n = R_n - S_n$  between each tap occurrence  $R_n$  and the corresponding stimulus occurrence  $S_n$  are recorded.  $T_n$  is the interstimulus interval, and  $e_n$  is called the synchronization error or asynchrony. (B) Typical time series and (C) histogram of the synchronization error  $e_n$  for an isochronous sequence (constant interstimulus interval of 500 ms; one trial from one subject). Note the normal tendency to anticipate the taps (called negative mean asynchrony, NMA, depicted as a dashed line around  $-20$  ms). The subject is normally unaware of his/her own NMA.

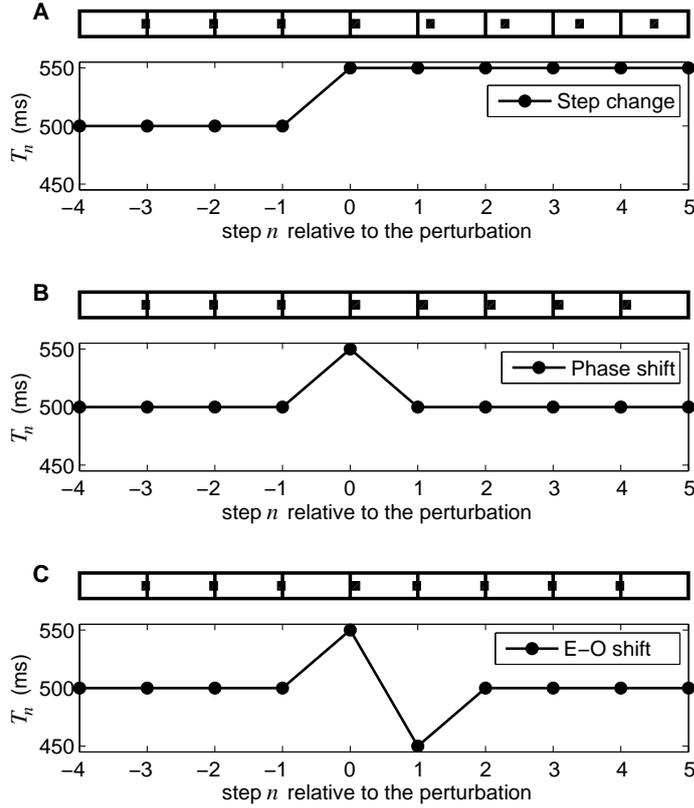


Figure 2: The three basic perturbation types. Interstimulus interval  $T_n$  as a function of position in the sequence, relative to the perturbation ( $n = 0$ ). (A) Step change: a single change in the interstimulus interval; it is considered the elementary perturbation in this work. (B) Phase shift: constant temporal shift of all tones following the perturbation, achieved by two consecutive elementary step changes. (C) Event-onset shift: shift of a single event, achieved by three consecutive elementary step changes. Top inset in every panel shows the corresponding tone onset times (solid squares) versus the non-perturbed period (thin vertical lines). After Repp (2005).

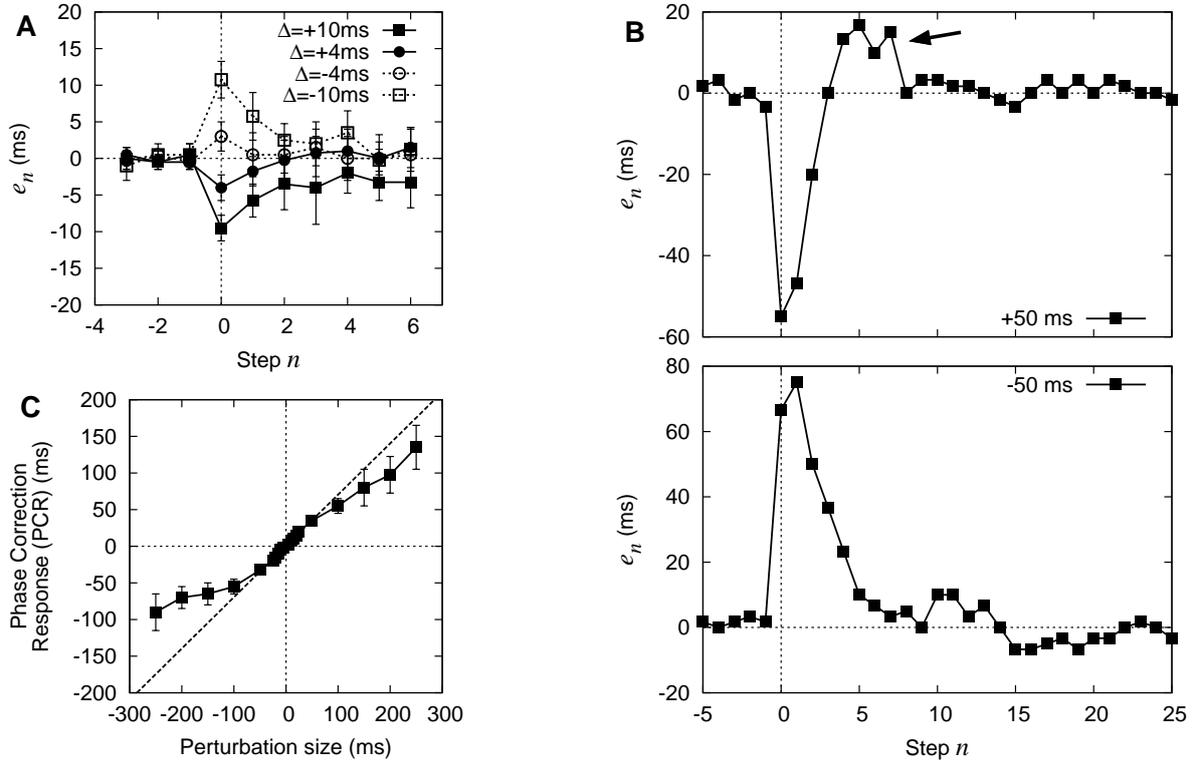


Figure 3: Review of published experimental results. (A) Phase-shift perturbations to a sequence with 500 ms pre-perturbation interstimulus interval. The error at step  $n = 0$  is a forced error, since the perturbation is unexpected and thus the subject is unaware of how large the perturbation will and when it will occur. Error correction is engaged already at the first step after the perturbation ( $n = 1$ ). The overall time evolution is a monotonic return to the baseline; the qualitative shape of the time evolution is not evidently dependent on the magnitude or the sign of the perturbation. Perturbations of  $\pm 10$  ms are near the detection threshold, while  $\pm 4$  ms are subliminal. Data digitized and re-plotted from Repp (2000) with permission. (B) Step-change perturbations to a sequence with 500 ms pre-perturbation interstimulus interval (*top*,  $\Delta = +50$  ms; *bottom*,  $\Delta = -50$  ms). Error correction is evident at the first step after the forced error. There is an asymmetric response to the perturbation: the return to the baseline is monotonic after the negative perturbation, but there is an overshoot (arrow in top panel) after the positive perturbation. Data digitized and re-plotted from Thaut et al. (1998) with permission. (C) Phase correction response (PCR), or relative shift of the tap immediately following a perturbation  $e_1 - e_0$ , for a phase-shift perturbation as a function of its magnitude, with a pre-perturbation interstimulus interval of 500 ms. The central portion of the function is linear, but the slope decreases for larger perturbations. This could be interpreted as a saturation effect in the subject's ability to correct the asynchrony after a perturbation of large magnitude. Data digitized and re-plotted from Repp (2002b) with permission.

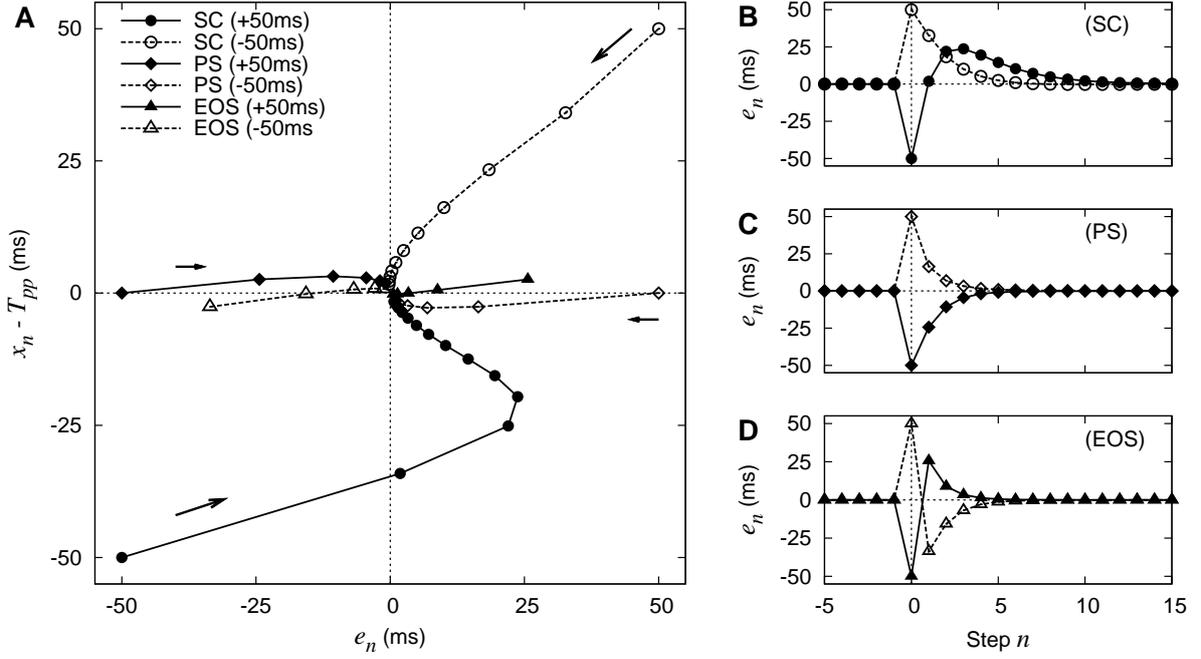


Figure 4: Model trajectories in phase space (A) and corresponding time series (B,C,D) (eq. 5). Six conditions are displayed: step changes (SC), phase shifts (PS), and event-onset shifts (EOS), each with two possible magnitudes  $\Delta = \pm 50\text{ms}$ . (A) Phase space. The trajectories are attracted to the stable fixed point (the point of subjective synchrony) after perturbing at  $n = 0$  an otherwise isochronous sequence with interstimulus interval  $T = 500$  ms. For visual clarity, only the portion of the trajectory after perturbation is plotted ( $n \geq 0$  for SC and PS;  $n \geq 1$  for EOS). (B) Step change: note that significant asymmetry in the time series only occurs in this perturbation. (C) Phase shift: synchrony is recovered after mostly symmetric, monotonic approach to baseline. (D) Event-onset shift: note the “switched” behavior: the positive (negative) perturbations are corrected from above (below).  $T_{pp}$  stands for “post-perturbation interstimulus interval”,  $T_{pp} = T + \Delta$ . Parameter values are displayed in Table 1.

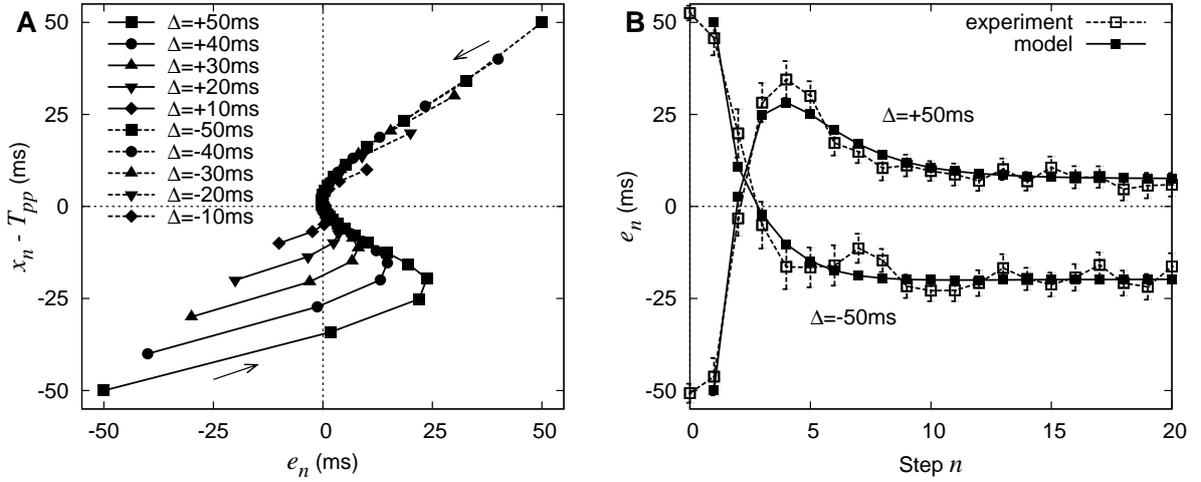


Figure 5: Model's prediction for step-change perturbations. (A) Model trajectories after a step-change perturbation of variable magnitude ( $\Delta = \pm 50, \pm 40, \pm 30, \pm 20$ , and  $\pm 10$  ms). Note the organization of the trajectories after the perturbation, particularly the overshoot in  $e_n$  after positive perturbations (i.e. from below): a prediction of our model is that the amount of overshoot would increase with the perturbation magnitude, while negative perturbations would not display overshoot. Parameter values are displayed in Table 1. (B) Fitting of the model to our experimental data (see Materials and Methods). Experimental data: mean across subjects  $\pm$  standard error of the mean. The model was fitted using the  $\pm 50$  ms step-change perturbations only (shown here; the complete dataset is shown in Figure 6).

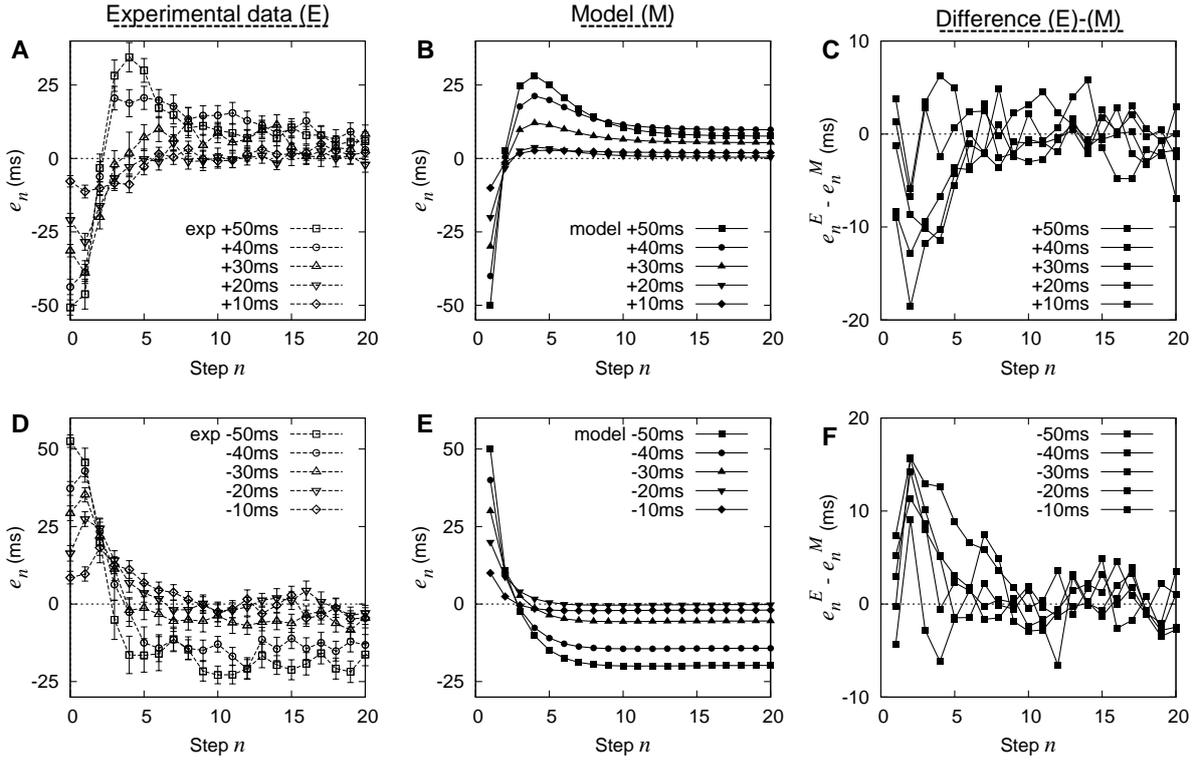


Figure 6: Experiment and simulation of positive (*top row*) and negative (*bottom row*) step-change perturbations. (*Left column*) Experimental results from our finger tapping task (responses averaged across subjects, with standard error bars based on the variability among subjects). As predicted by our model, note in panel (A) the increasing overshoot with increasing perturbation magnitude. (*Middle column*) Time series of the model. All model simulations throughout this work use the same set of parameter values (see Table 1). (*Right column*) Difference between the corresponding experimental (E) and model (M) time series. The difference is greater at smaller perturbation magnitudes, due to the experimental series developing the overshoot more slowly than the model series.

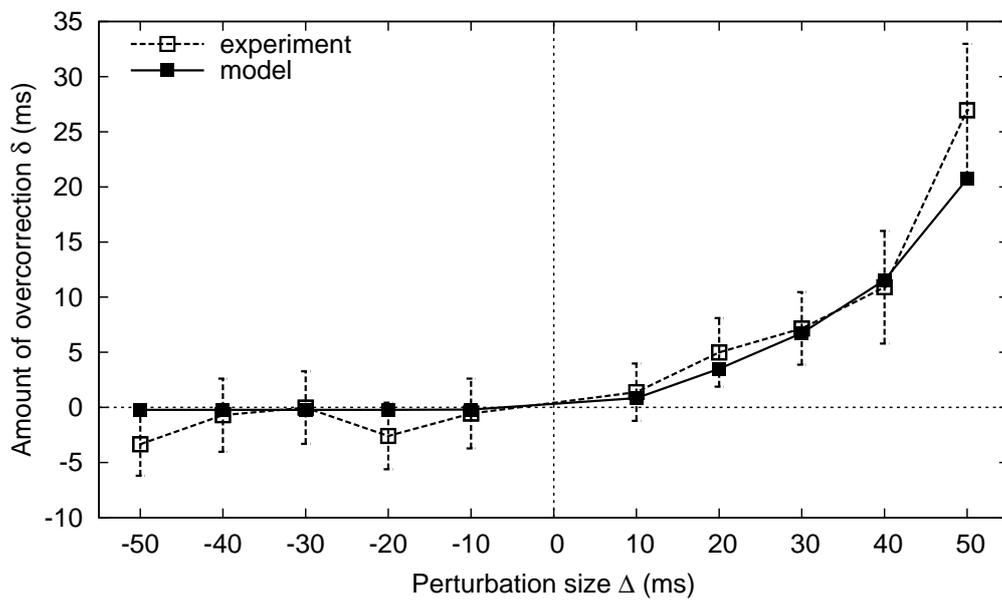


Figure 7: Direct comparison between our model’s prediction for step-change perturbations and our experimental data. The amount of overcorrection  $\delta$  is zero for negative perturbations, whereas for positive perturbations it is a nonlinear, increasing function of the perturbation size. Each point here was computed from the corresponding time series in Figure 6, thus the only points coming from fitting are the two extremes  $\Delta = \pm 50$  ms.

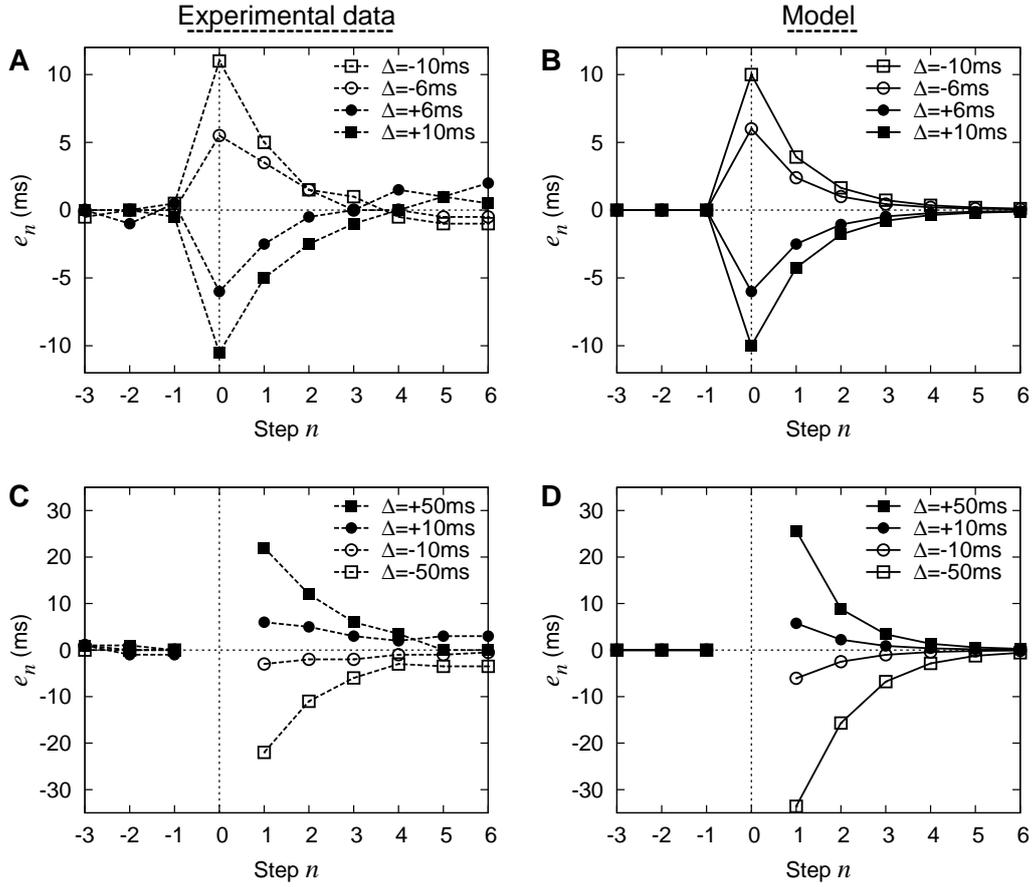


Figure 8: Predictions for other perturbation types. Published experimental data (*left*) and numerical simulations from our model (*right*), for phase-shift (*top*) and event-onset shift (*bottom*) perturbations. The similarity is remarkable, especially because none of these perturbations were included in the building of the model or used for the fitting. All simulations in this work were performed with the same set of parameter values (Table 1). Experimental phase shifts: data digitized and re-plotted from Repp (2001a); experimental event-onset shifts: data digitized and re-plotted from Repp (2002a); all with permission.

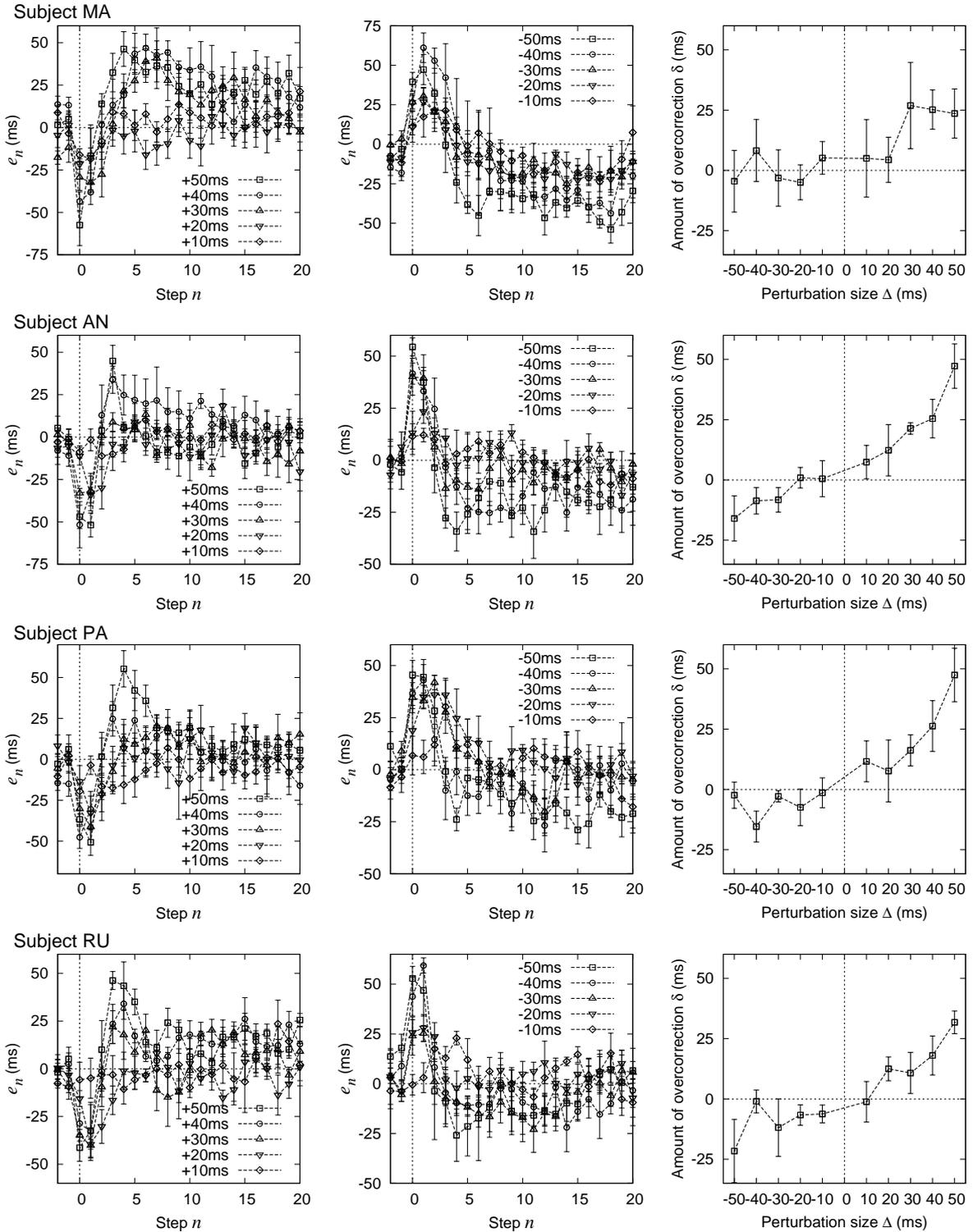


Figure 9: Individual data from four subjects. (Left) Positive and (Middle) negative perturbations; mean of 5 trials  $\pm$  standard error for each perturbation size. Despite interindividual differences (like the slower overshoot and return to baseline of subject PA), the qualitative behavior is common to all subjects: overshoot in the left column but not in the right column. (Right) Amount of overcorrection  $\delta$ . Note the greater values of overcorrection for positive perturbations.

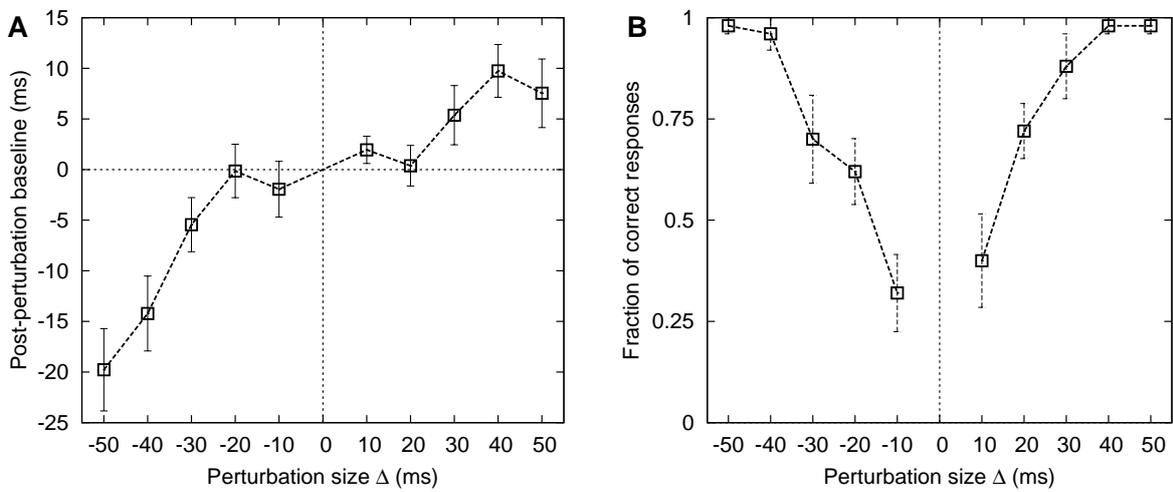


Figure 10: (A) Post-perturbation baseline (relative to its pre-perturbation value). The baseline after a perturbation increases as a function of perturbation size (i.e. it becomes more positive for slower sequences), which cannot be explained by the known decreasing dependence of NMA on interstimulus interval for auditory sequences. Mean across subjects  $\pm$  standard error of the mean. (B) Detection of perturbations: average percentage of trials where the perturbation was detected. The  $\pm 10$  ms perturbations are just below the detection threshold (0.5). Mean across subjects  $\pm$  standard error of the mean.