Two-Phase Tempo Increase in Collective Clapping: Acceleration–Plateau Dynamics and Modality Effects

Yuka Koike (koike.yuka.qp@alumni.tsukuba.ac.jp)[1] & Takayuki Niizato[1]
[1]Department of Intelligent Interaction Technologies, Institute of Systems and Information Engineering, University of Tsukuba

Abstract

In group rhythmic behavior, the tempo increase phenomenon can emerge even when participants appear well synchronized. We analyzed collective clapping from 49 participants while manipulating modality (auditory/visual/audiovisual) and group size (pair/trio). We tested a sequential effect in which beat-wise asynchrony $A(k)$ predicts the subsequent tempo change $\Delta\bar{\mu}(k)$. For each trial, we estimated accel and plateau phases by fitting a one-breakpoint piecewise linear regression to the smoothed tempo series and selecting the breakpoint k_p via an AIC-based grid search, using only the tempo series to avoid circular inference. In the confirmatory analysis ($w = 7$), an AR(1) mixed-effects model revealed a robust sequential effect: larger asynchrony predicted a more negative immediate tempo change (shorter subsequent ICI). The coupling was stronger during accel and attenuated toward zero during plateau, and it was strongest in the visual condition during accel. An exploratory check across $w = 1 \ldots 15$ yielded consistent patterns.

Keywords: Collective behavior, tempo increase phenomenon, sequential effect, modality

Introduction

In group rhythmic behavior, participants can appear tightly synchronized even as the collective tempo gradually increases across a trial. The tempo increase phenomenon is particularly evident in collective clapping: even when individual beats align closely, the group-level tempo increases as the sequence unfolds, an apparently paradoxical pattern (Thomson et al., 2018). The tempo increase phenomenon is more likely to emerge under social interaction than during solo performance (Knoblich et al., 2011; Okano et al., 2017), and classic work on audience clapping showed that clapping frequency can continue to increase even after the group reaches unison (Néda et al., 2000). A dominant account explains this phenomenon within an error-correction framework, including phase correction and period correction, whereby performers adjust the timing of upcoming actions in response to observed synchronization errors (Repp & Su, 2013; Wolf & Knoblich, 2022). At the same time, prior work argues that this phenomenon may reflect additional mechanisms beyond a single correction process, such as systematic forward shifts in action timing (phase advance) and interaction-driven changes in internal timekeeping during joint action (Wolf & Knoblich, 2022; Wolf et al., 2019).

Despite this progress, two key issues remain unresolved. First, it is unclear whether the tempo increase phenomenon progresses approximately uniformly throughout a trial or instead emerges early and then plateaus as coordination consolidates (Thomson et al., 2018). Second, existing characterizations often rely on aggregate indices of synchrony, which can obscure the mechanisms that generate the tempo increase phenomenon (Keller et al., 2014; Repp & Su, 2013). Collective clapping exhibits qualitative transitions in coordination (e.g., self-organized shifts into synchrony), suggesting that becoming synchronized and subsequent tempo transformation may reflect distinct within-trial regimes (Néda et al., 2000). We therefore focus on a mechanistic, beat-to-beat question: to what extent does a deviation at one beat drive the immediate tempo change that follows? Specifically, we test whether inter-participant asynchrony at beat k, $A(k)$, predicts the subsequent tempo change, $\Delta\bar{\mu}(k)$ (a sequential effect), and whether this coupling varies across the accel and plateau phases and across sensory modalities. We quantify the sequential effect linking asynchrony to immediate tempo adjustment in collective clapping and compare its strength across (i) within-trial phases (accel vs. plateau) and (ii) experimental conditions (modality: auditory/visual/audiovisual; group size: pair/trio) (Dotov et al., 2022). Rather than modeling the tempo increase phenomenon with a single constant slope, we partition each trial's tempo trajectory into two phases in a data-driven way. We fit a piecewise linear regression with one breakpoint to the smoothed tempo series, estimate the breakpoint k_p, and label beats as accel for $k \leq k_p$ and plateau for $k > k_p$. Crucially, we estimate phase boundaries using only the tempo series and never use $A(k)$ for segmentation, preventing circular inference. This design enables a within-trial test of whether the sequential effect differs between the accel phase and the plateau phase. Our central hypothesis posits a sequential effect: larger deviations $A(k)$ induce larger immediate tempo adjustments, yielding shorter subsequent ICIs and thus more negative $\Delta\bar{\mu}(k)$ (Okano et al., 2019). We further predict that this coupling is phase dependent, with a stronger sequential effect during the accel phase and attenuation toward zero during the plateau phase. We also test for modality dependence (auditory, visual, audiovisual), because the availability and reliability of temporal information differ across sensory channels (Repp & Penel, 2004; Repp & Su, 2013). Prior work suggests modality-specific constraints on temporal precision, motivating our comparison of conditions in which timing information is conveyed through different sensory channels. Specifically, the tempo increase phenomenon

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

may reflect a modality-general interaction mechanism, or instead depend on which sensory channel conveys timing information. Methodologically, we compute $\Delta\bar{\mu}(k)$ using a trailing average that does not incorporate future information. We then estimate the sequential relationship $A(k) \rightarrow \Delta\bar{\mu}(k)$, and its modulation by phase and condition, using a mixed-effects model with AR(1) residuals to account for within-trial autocorrelation. This study makes three contributions. First, we operationalize the tempo increase phenomenon not as an aggregate property of mean synchrony but as a beat-to-beat sequential effect, namely how strongly a deviation at one beat drives the subsequent tempo change, providing a more direct measure of error-correction dynamics (Wolf & Knoblich, 2022). Second, we characterize within-trial dynamics with a two-phase model (accel/plateau) based on breakpoint estimation, without assuming a uniform tempo increase across the trial. Third, by manipulating modality and group size, we compare the sequential effect across conditions and test for mechanistic differences, including the contribution of visual timing information. Together, our framework provides a basis for understanding how groups adjust timing from beat to beat as a function of within-trial phase and sensory context.

Methods

Participants

Forty-nine volunteers participated in the study (14 women, 35 men; mean age = 22.3 years, SD = 1.8).

Design

We manipulated modality (auditory/visual/audiovisual) and group size (pair/trio) as between-subject factors, with each participant assigned to a single condition. Participants completed 7–8 trials, and we collected 20 trials per condition.

Procedure

Each trial was a synchronization–continuation clapping task designed to yield 225 claps per participant. Participants first synchronized their clapping to a metronome at 75 bpm and then continued clapping after the metronome ceased. We instructed participants to match one another's clap timing as closely as possible throughout the trial. The modality manipulation controlled what information about others' claps was available (auditory, visual, or audiovisual). To prevent the results from depending on a fixed pair or trio, we reassigned group membership across trials.

Group tempo: $\mu(k)$ We define each participant's inter-clap interval (ICI) as

$$\mathrm{ICI}_i(k) = T_i(k+1) - T_i(k),$$

and we define the group-level tempo at beat k as the participant-average ICI:

$$\mu(k) = \frac{1}{N}\sum_i \mathrm{ICI}_i(k),$$

where N is the group size in that trial (2 or 3).

Trailing moving average $\bar{\mu}_w(k)$ and tempo change $\Delta\bar{\mu}_w(k)$ To align the analysis with our sequential hypothesis and to avoid contaminating the predictor–outcome relationship with future information, we smooth the tempo series using a trailing (backward-looking) moving average. For a window width w, we define

$$\bar{\mu}_w(k) = \frac{1}{w}\sum_{r=0}^{w-1} \mu(k-r) \quad (k \geq w).$$

We define the immediate tempo change as the first difference of the smoothed series:

$$\Delta\bar{\mu}_w(k) = \bar{\mu}_w(k+1) - \bar{\mu}_w(k).$$

The primary analyses use $w = 7$. Note that $\Delta\bar{\mu}_w(k) < 0$ indicates a shorter subsequent ICI, i.e., a change toward acceleration.

Robustness analysis (generalization across w). As a robustness check, we also compute $(\bar{\mu}_w, \Delta\bar{\mu}_w)$ for $w = 1, \ldots, 15$ and visualize how the estimated effect sizes vary with the smoothing window. However, confirmatory statistical inference (e.g., p-value-based conclusions) is restricted to $w = 7$.

Preprocessing: clap times and inter-clap intervals

We extracted clap times as peak locations $T_i(k)$ from each participant's composite acceleration signal (Fig. 1), where $k = 1, 2, \ldots$ indexes beats. We computed inter-clap intervals as

$$\mathrm{ICI}_i(k) = T_i(k+1) - T_i(k).$$

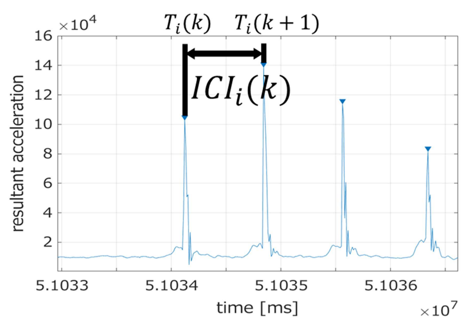

Figure 1: Example peak detection from the composite clap acceleration signal. Peaks at times $T_i(k)$ and $T_i(k+1)$ define the inter-clap interval $\mathrm{ICI}_i(k) = T_i(k+1) - T_i(k)$.

Measures and definitions

Asynchrony: $A(k)$ We define the inter-participant timing deviation at beat k as the mean absolute pairwise difference in clap times:

$$A(k) = \mathrm{mean}_{(i,j)} \left|T_i(k) - T_j(k)\right|.$$

We treat $A(k)$ as a continuous measure of synchrony and do not binarize it using a threshold in the primary analyses.

Given the estimated breakpoint k_p, we label each beat k as

- **accel** if $k \leq k_p$,
- **plateau** if $k > k_p$,

and we use this phase label in all subsequent analyses (estimation of sequential effects and condition comparisons).

A critical design feature is that phase estimation relies only on the tempo series $\bar{\mu}_w(k)$. We do not use the asynchrony measure $A(k)$ (the predictor in the sequential-effect model) to determine k_p. This separation prevents circular inference ("double dipping"), in which one would define phases using $A(k)$ and then claim phase differences in the effect of $A(k)$.

Temporal alignment of predictor and outcome

Our central hypothesis is that asynchrony at beat k drives the immediate tempo change that follows. Accordingly, we align the predictor and outcome as follows:

- **Predictor:** beat-k asynchrony $A(k)$ (or, when required, a smoothed variant $\bar{A}(k)$).
- **Outcome:** tempo change from beat k to beat $k+1$,

$$\Delta\bar{\mu}_w(k) = \bar{\mu}_w(k+1) - \bar{\mu}_w(k).$$

This alignment yields a direct interpretation of the sequential effect $A(k) \rightarrow \Delta\bar{\mu}_w(k)$: the model predicts the next-interval tempo adjustment after observing the deviation at beat k. Because $\bar{\mu}_w(k)$ is a trailing average, $\Delta\bar{\mu}_w(k)$ depends only on information available up to beat $k+1$ and does not reference future beats.

Phase estimation: piecewise regression with AIC-based grid search (without using $A(k)$)

Because the tempo increase phenomenon does not necessarily progress uniformly across a trial, we segment each trial's tempo trajectory into two within-trial phases: accel and plateau. We estimate the phase boundary separately for each trial rather than imposing a fixed split.

Specifically, we fit a piecewise linear regression with a single breakpoint to the smoothed tempo series $\bar{\mu}_w(k)$ computed with the primary window width ($w = 7$). Let k_p denote the breakpoint. To control degrees of freedom and avoid unstable boundary estimates, we restrict k_p to a discrete candidate set and select the breakpoint that minimizes AIC. We prespecify the candidate set to correspond to 30/45/60/75/90 intervals and choose the AIC-minimizing k_p within this range.

Autocorrelation: AR(1) residual process Within a trial, $\Delta\bar{\mu}_w(k)$ evolves continuously over time, and residuals from adjacent beats are likely correlated. We model the residual term $\varepsilon(k)$ as an AR(1) process,

$$\varepsilon(k) = \rho\,\varepsilon(k-1) + u(k), \quad |\rho| < 1,$$

thereby explicitly capturing serial dependence within trials and improving the validity of standard errors and statistical tests for the fixed effects.

Fixed effects and model comparison For the primary analysis ($w = 7$), we compared a baseline model (m0; phase $\times$ $\bar{A}$ + modality $\times$ group_size) to an extended model allowing modality to modulate the sequential effect (m1; phase $\times$ $\bar{A}$ $\times$ modality + modality $\times$ group_size). Both models used the same random-effects structure (random intercept for trial) and the same AR(1) residual correlation within trial, and were compared using a likelihood-ratio test under ML estimation.

Inference strategy and multiple-comparison control (confirmatory vs. exploratory) We restrict confirmatory inference (i.e., p-value-based conclusions) to the pre-specified primary analysis using $w = 7$. As a robustness check, we additionally estimate the same models for window widths $w = 1, \ldots, 15$ and visualize how the effect size and direction vary as a function of w. We present these results as exploratory diagnostics of stability and do not use them to claim statistical significance.

Statistical model: mixed-effects model with AR(1) errors

Sequential effects are estimated from beat-by-beat time series, so observations within each trial are not independent. We therefore fit a mixed-effects model with an AR(1) error structure to account for within-trial dependence.

Outcome and predictors

- **Outcome:** immediate tempo change in the next interval, $\Delta\bar{\mu}_w(k)$ (primary analyses use $w = 7$).
- **Key predictor:** asynchrony at beat k (e.g., a smoothed version $\bar{A}(k)$ when appropriate).
- **Moderators / comparison factors:** phase (accel/plateau), modality (auditory/visual/audiovisual), group size (pair/trio), and their interactions.

The focal parameter is the coefficient of $\bar{A}(k)$ in the regression of $\Delta\bar{\mu}_w(k)$ (the sequential effect). We test phase dependence via the $\bar{A}\times$phase interaction. We test modulation by modality and group size via the corresponding interaction terms, including a three-way interaction when necessary to capture condition differences in phase-dependent sequential effects.

Random effects (hierarchical structure) We include trial as a random effect to account for repeated measurements and trial-to-trial baseline differences (e.g., globally faster or slower trials). Concretely, we fit a random intercept for trial, allowing us to estimate the sequential effect while separating it from between-trial variability.

Results

Confirmatory analyses ($w = 7$): sequential effect, phase dependence, and modality dependence

In the confirmatory analyses, we used the tempo-change outcome $\Delta\bar{\mu}_w(k)$ computed from a trailing moving average with window width $w = 7$. We tested whether asynchrony at beat k (e.g., $\bar{A}(k)$) predicts the immediate tempo adjustment in the subsequent interval. We estimated this relationship using a mixed-effects model with an AR(1) residual structure to account for within-trial serial dependence.

We first compared a baseline model including the phase × $\bar{A}$ interaction (m0) with an extended model including the phase × $\bar{A}$ × modality interaction (m1). The extended model significantly improved model fit, $\chi^2(6) = 56.63$, $p < .001$, indicating that the sequential effect varied by modality. We therefore report the extended model as the primary confirmatory model. Focal fixed-effect estimates from this model are shown in Table 1.

Table 1: Focal fixed effects from the confirmatory AR(1) mixed-effects model.

Term	Est.	SE	95% CI	p
$\bar{A}$	-0.083	0.020	[-0.123, -0.043]	< .001
plateau ×$\bar{A}$	0.093	0.022	[0.050, 0.137]	< .001
$\bar{A}$× visual	-0.324	0.066	[-0.453, -0.196]	< .001
$\bar{A}$× audiovisual	-0.003	0.031	[-0.065, 0.059]	.922
plateau ×$\bar{A}$× visual	0.275	0.070	[0.137, 0.412]	< .001
plateau ×$\bar{A}$× audiovisual	0.000	0.034	[-0.066, 0.067]	.994

Reference levels were audio, accel, and pair. The outcome was $\Delta\bar{\mu}_w(k)$ with $w = 7$. $\bar{A}$ denotes smoothed beat-wise asynchrony.

The model showed a reliable sequential effect in the reference condition (audio, accel): larger $\bar{A}$ predicted a more negative subsequent tempo change, corresponding to a shorter subsequent ICI and therefore an acceleration-direction adjustment (Table 1). The positive phase plateau × $\bar{A}$ interaction indicated that this coupling was attenuated in the plateau phase. The modality interactions further showed that the accel-phase coupling was stronger in the visual condition than in the audio condition, whereas the audiovisual condition did not differ from the audio condition in the accel phase.

To make these interaction effects easier to interpret, we computed condition-specific simple slopes of $\bar{A}$ for each modality × phase combination (Table 2). During the accel phase, $\bar{A}$ significantly predicted negative tempo change in all modality conditions, with the strongest effect in the visual condition. During the plateau phase, the simple slopes were small and did not differ significantly from zero in any modality condition.

Table 2: Simple slopes of $\bar{A}$ by modality and phase.

Modality	Phase	Slope	SE	95% CI	p
audio	accel	-0.083	0.020	[-0.123, -0.043]	< .001
visual	accel	-0.407	0.062	[-0.529, -0.285]	< .001
audiovisual	accel	-0.086	0.024	[-0.133, -0.039]	< .001
audio	plateau	0.011	0.009	[-0.007, 0.028]	.238
visual	plateau	-0.039	0.026	[-0.091, 0.013]	.138
audiovisual	plateau	0.008	0.010	[-0.011, 0.027]	.418

Slopes represent the estimated effect of $\bar{A}$ on $\Delta\bar{\mu}_w(k)$ within each modality × phase cell. Negative slopes indicate acceleration-direction coupling.

The same condition-specific simple slopes are visualized in Figure 2.

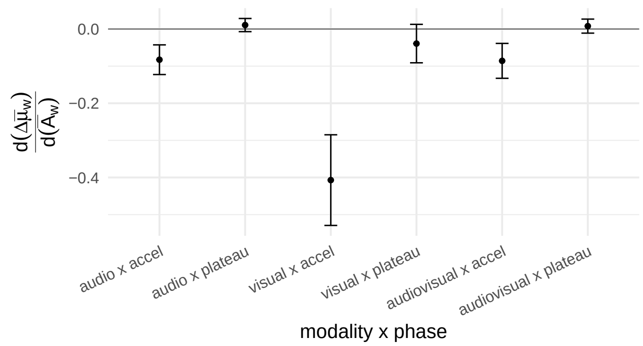

Figure 2: Main analysis ($w = 7$): condition-specific simple slopes of $d(\Delta\bar{\mu}_w)/d\bar{A}_w$ across modality × phase. Error bars show 95% CIs computed from fixed-effect linear contrasts; values below zero indicate acceleration-direction coupling (larger $\bar{A}_w$ predicts more negative $\Delta\bar{\mu}_w$), whereas values above zero indicate deceleration-direction coupling.

Exploratory robustness ($w = 1 \ldots 15$): effect-size curves

Although confirmatory inference was restricted to the pre-specified window width ($w = 7$), we assessed robustness to the choice of smoothing scale by fitting the same model across window widths $w = 1, \ldots, 15$ and visualizing the asynchrony coefficient (the sequential-effect slope) as an effect-size curve. This analysis was exploratory and was not used to make claims about statistical significance.

Across a broad range of window widths, the qualitative pattern remained stable: the sequential effect retained a negative sign (acceleration direction) and tended to be stronger in the accel phase and weaker in the plateau phase. Thus, varying the degree of temporal smoothing did not materially change the conclusion that beat-wise deviations predict immediate tempo adjustments and that this coupling is phase dependent. Modality-related tendencies, in particular the larger sequential effect in the visual condition during the accel phase, also appeared in the same direction for many values of w.

Taken together, these robustness checks suggest that the sequential effect, its phase dependence, and its modality dependence in the primary analysis ($w = 7$) are unlikely to be artifacts of a particular smoothing choice and instead reflect a pattern that is stable across multiple temporal scales.

Figure 3 shows how key coefficient estimates vary as a function of window width w (dashed line: $w = 7$).

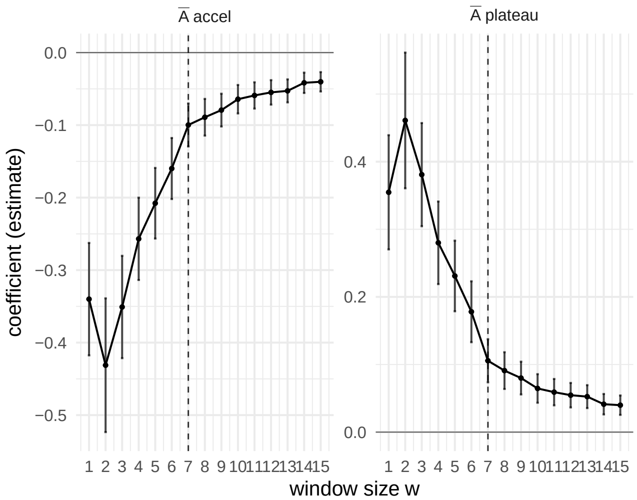

Figure 3: Exploratory robustness check: coefficient curves across window widths $w = 1 \ldots 15$ (dashed line: $w = 7$). The figure summarizes effect sizes across smoothing scales rather than supporting significance claims.

Discussion

This study reframed the tempo increase phenomenon in collective clapping not as an aggregate index of "how well the group synchronizes on average," but as a mechanistic sequential effect (Wolf & Knoblich, 2022; Wolf et al., 2019): how strongly a deviation at one beat drives the subsequent tempo change. We found that larger asynchrony $A(k)$ predicted a more negative subsequent tempo change $\Delta\bar{\mu}(k)$ (i.e., a shorter ICI), indicating beat-to-beat error correction (Repp & Su, 2013). This pattern suggests that the tempo increase phenomenon in collective clapping is not simply the accumulation of random fluctuations. Instead, it reflects the cumulative consequence of repeated local adjustments in an error→correction loop. Importantly, this sequential relationship remained robust when we modeled within-trial serial dependence using a mixed-effects model with AR(1) errors, supporting the interpretation that asynchrony predicts subsequent tempo change even after accounting for autocorrelation.

Second, we did not assume that the tempo increase phenomenon progresses uniformly across a trial. Instead, we segmented each trial into accel and plateau phases by estimating a breakpoint on the tempo trajectory and compared the sequential effect across phases (Yozevitch et al., 2023; Zamm et al., 2023). The results revealed clear phase dependence: the sequential effect was strong during the accel phase and attenuated toward zero during the plateau phase (Heggli et al., 2021). This dissociation suggests two within-trial regimes: an early regime in which deviations trigger strong immediate adjustments that collectively increase tempo, and a later regime in which coordination tightens and deviations are less likely to translate into systematic tempo change. Put differently, the gain of error correction appears to decrease within a trial, as the strength with which deviations are converted into tempo change diminishes over time. This phase dependence addresses the ambiguity of whether the tempo increase phenomenon proceeds steadily or instead emerges early and then plateaus by grounding the answer in a mechanistic index of sequential coupling.

Third, the magnitude of the sequential effect depended on sensory modality, with the strongest slopes in the visual condition during the accel phase (Repp & Penel, 2004). This pattern may seem to contradict the widely reported auditory dominance in timing and synchronization (Repp & Su, 2013). However, our sequential effect does not measure synchrony precision per se. Instead, it measures how strongly deviations are translated into immediate tempo adjustments, that is, the correction gain. Under visual-only information, temporal cues to clap timing are comparatively coarse and intermittent, which can make moment-to-moment alignment harder to maintain. In this setting, when a deviation becomes apparent, participants may apply larger corrective updates, yielding stronger deviation-to-tempo coupling. By contrast, auditory and audiovisual conditions provide richer timing cues that support finer-grained adjustments, which may reduce the need for large tempo shifts and yield weaker coupling overall and near-zero coupling in the plateau phase. Thus, the modality difference is not a ranking of which condition yields the best synchrony but evidence that the rule converting deviations into tempo change differs across sensory contexts.

In summary, our results indicate that (i) collective clapping exhibits a sequential effect in which asynchrony predicts subsequent tempo change, (ii) the strength of this coupling declines from the accel phase to the plateau phase within trials, and (iii) the sequential effect is modulated by modality, with particularly large effects in the visual accel phase. By reframing the tempo increase phenomenon as a mechanistic quantity, namely the gain linking deviations to tempo change, our framework enables interpretable comparisons across phases and experimental conditions and clarifies how interaction dynamics reshape timing control during collective synchronization.

Future work

Direction-sensitive measures (lead/lag) and separating phase vs. period components. Our current asynchrony measure $A(k)$ is defined as the mean absolute pairwise time difference across participant pairs and therefore discards directional information about who leads and who lags. A natural next step is to introduce signed phase differences (lead/lag) and test whether the sequential effect reflects distinct interaction modes, such as "recovery from delay" (laggers catching up) versus "following a leader" (others adapting to a consistently leading participant). In addition, the tempo-change outcome $\Delta\bar{\mu}(k)$ may conflate multiple correction processes, reflecting both within-cycle phase correction and longer-term period correction. Future work should therefore adopt measures and models that jointly represent phase and period dynamics, allowing us to decompose the sequential effect into phase-driven versus period-driven components. This decomposition is particularly important for interpreting the strong sequential effect in the visual condition because it can distinguish whether visual rushing primarily reflects increased phase correction, stronger period updates, or a combination of the two.

Modeling individual differences via random slopes. The sequential effect $A(k) \to \Delta\bar{\mu}(k)$, interpreted as the gain of error correction, may vary across individuals and contexts. Future analyses can extend the current AR(1) mixed-effects framework by adding participant-level random slopes for the sequential predictor and, where appropriate, allowing these random slopes to vary by experimental condition. This extension would allow us to estimate how strongly each participant adjusts tempo across conditions, providing a more precise characterization of inter-individual variability in error-driven tempo control during joint rhythmic action.

Together, these directions move beyond the present aggregate characterization and sharpen the mechanism underlying the tempo increase phenomenon by incorporating directionality (lead/lag), component-level decomposition (phase vs. period), and person-specific variability in correction gain.

Acknowledgments

We used ChatGPT to generate improvement suggestions for parts of the analysis code (preprocessing, visualization, and model implementation) to support debugging and refactoring. The authors reviewed the generated content iteratively and adopted it only after verifying reproducibility on known test cases. The study design, statistical decisions, and conclusions were made by the authors.

References

Dotov, D., Delasanta, L., Cameron, D. J., Large, E. W., & Trainor, L. (2022). Collective dynamics support group drumming, reduce variability, and stabilize tempo drift. *Elife*, *11*, e74816.

Heggli, O. A., Konvalinka, I., Kringelbach, M. L., & Vuust, P. (2021). A metastable attractor model of self–other integration (meamso) in rhythmic synchronization. *Philosophical Transactions of the Royal Society B*, *376*(1835), 20200332.

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658).

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. *Psychology of learning and motivation*, *54*, 59–101.

Néda, Z., Ravasz, E., Vicsek, T., Brechet, Y., & Barabási, A.-L. (2000). Physics of the rhythmic applause. *Physical Review E*, *61*(6), 6987.

Okano, M., Kurebayashi, W., Shinya, M., & Kudo, K. (2019). Hybrid dynamics in a paired rhythmic synchronization–continuation task. *Physica A: Statistical Mechanics and Its Applications*, *524*, 625–638.

Okano, M., Shinya, M., & Kudo, K. (2017). Paired synchronous rhythmic finger tapping without an external timing cue shows greater speed increases relative to those for solo tapping. *Scientific Reports*, *7*(1), 43987.

Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological research*, *68*(4), 252–270.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic bulletin & review*, *20*(3), 403–452.

Thomson, M., Murphy, K., & Lukeman, R. (2018). Groups clapping in unison undergo size-dependent error-induced frequency increase. *Scientific reports*, *8*(1), 1–9.

Wolf, T., & Knoblich, G. (2022). Joint rushing alters internal timekeeping in non-musicians and musicians. *Scientific Reports*, *12*(1), 1190.

Wolf, T., Vesper, C., Sebanz, N., Keller, P. E., & Knoblich, G. (2019). Combining phase advancement and period correction explains rushing during joint rhythmic activities. *Scientific Reports*, *9*(1), 9350.

Yozevitch, R., Dahan, A., Seada, T., Appel, D., & Gvirts, H. (2023). Classifying interpersonal synchronization states using a data-driven approach: Implications for social interaction understanding. *Scientific Reports*, *13*(1), 11150.

Zamm, A., Debener, S., & Sebanz, N. (2023). The spontaneous emergence of rhythmic coordination in turn taking. *Scientific Reports*, *13*(1), 3259.