The Timing of Regular Sequences: Production, Perception, and Covariation

Michael Schwartze[1] and Sonja A. Kotz[1,2]

Abstract

■ The temporal structure of behavior provides information that allows the tracking of temporal regularity in the sensory and sensorimotor domains. In turn, temporal regularity allows the generation of predictions about upcoming events and to adjust behavior accordingly. These mechanisms are essential to ensure behavior beyond the level of mere reaction. However, efficient temporal processing is required to establish adequate internal representations of temporal structure. The current study used two simple paradigms, namely, finger-tapping at a regular self-chosen rate (spontaneous motor tempo) and ERPs of the EEG (EEG/ERP) recorded during attentive listening to temporally regular and irregular "oddball" sequences to explore the capacity to encode and use temporal regularity in production and perception. The results show that specific aspects of the ability to time a regular sequence of events in production covary with the ability to time a regular sequence in perception, probably pointing toward the engagement of domain-general mechanisms. ■

INTRODUCTION

The ability to use the temporal structure of past events to predict when a future event will occur is a remarkable capacity of neurocognitive function. Such temporal extrapolation requires an online analysis of temporal relations and the extraction of regularities from a sequence of events (Bendixen, SanMiguel, & Schröger, 2012; Winkler, Denham, & Nelken, 2009). Temporal extrapolation provides the opportunity to optimize behavior by means of predictive adaptation, which is indispensable for the efficient coping with a dynamically changing environment. Predictive adaptation has many facets and may refer to widely divergent anticipatory mechanisms, for example, a change in the orientation of the head to improve sound perception, the selection of a particular action to improve response speed, or the allocation of attention to specific features of an input to improve the processing of critical information (Nobre, Correa, & Coull, 2007). Together, these mechanisms may bias to be appropriately ahead of time to optimally respond to particular events. However, the merit of predictive adaptation is tightly coupled with an individual's ability to infer regular relations among events to resolve two issues: what type of event is to be expected (formal regularity) and when (temporal regularity) is it to be expected (Costa-Faidella, Baldeweg, Grimm, & Escera, 2011; Schwartze, Rothermich, Schmidt-Kassow, & Kotz, 2011)? The behavioral advantage of not only knowing what type of event will happen but also knowing precisely when an event will occur is one of the main arguments in favor of neural mechanisms that engage in temporal processing and in the recognition of temporal regularity in perception and production.

Several brain structures that play a key role in motor behavior are considered to form the neural substrate of dedicated temporal processing, namely, the cerebellum, the BG, and the SMA (Merchant, Harrington, & Meck, 2013; Coull, Cheng, & Meck, 2010). The fact that classical motor structures engage in dedicated temporal processing allows speculating about which common mechanisms underlie the ability to time events in production and perception. Dedicated temporal processing may factor into the basic predictive power of the motor system (Schubotz, 2007; Schütz-Bosbach & Prinz, 2007). However, the ability to generate precise temporal predictions may complement and partly determine the quality of higher level cognitive processes. For example, temporal prediction may support cross-modal temporal integration in the pFC (Fuster, 2001; Fuster, Bodner, & Kroger, 2000). However, evidence for shared timing mechanisms in production and perception is mixed. There are reports of such mechanisms (Ivry & Hazeltine, 1995), of a potential overlap in core structures such as the BG and the cerebellum in contrast to differential cortical activation (Bueti, Walsh, Frith, & Rees, 2008; Lewis & Miall, 2003), and of larger variability in perceptual than in motor tasks despite correlations between the two (Merchant, Zarco, & Prado, 2008). A meta-analysis of neuroimaging studies identified dissociable neural correlates and also a domain-general "core network," which comprises supplementary motor and prefrontal areas (Wiener, Turkeltaub, & Coslett, 2010). However, even if timing in production and perception

[1]University of Manchester, [2]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

relies on different neural mechanisms, these mechanisms may be inseparable in the motor domain. Motor behavior implies circular processing within a sensorimotor perception–action cycle, in which actions produce changes in the environment, leading to new sensory inputs, which lead to new actions (Fuster, 2001). Consequently, it is inherently difficult to isolate perceptual aspects that contribute to the ability to time a regular sequence of events from tasks that involve a motor component, irrespective of the fundamental dissociation of central timing and peripheral motor processes (Wing & Kristofferson, 1973).

The most simple and perhaps most naturalistic experimental setup to explore the timing of regular sequences in the sensorimotor domain would require participants to produce a regular sequence of repetitive actions in the absence of extroceptive stimulation. Corresponding measures of unpaced motor performance have been discussed in terms of internal tempo, personal tempo, mental tempo, or spontaneous motor tempo (SMT; for a review, see Fraisse, 1982). SMT has been described as independent of simple biomechanical mechanisms and centered at around 600 msec within a range from 400 to 900 msec (Fraisse, 1982). However, SMT seems to slow down across the lifespan so that a mean of 500 msec and a range from about 300 to 1100 msec are probably more representative across different age groups (MacDougall & Moore, 2005; Moelants, 2002; Vanneste, Pouthas, & Wearden, 2001). Individual SMT rates have been found to be highly correlated with a perceptual counterpart termed the "preferred perceptual tempo" (McAuley, Jones, Holub, Johnston, & Miller, 2006).

In the current study, two runs of SMT data were recorded from a participant group with a broad age range to assess the efficient production of temporal regularity in a sensorimotor task. In line with previous findings, we expected SMT to provide a relatively stable and reproducible measure of individual performance. Primary variables of interest were rate and variability, which were considered as global measures of task performance and which were expected to be similar across the two runs. In addition to these global measures, a number of additional variables were extracted from the first run. These variables related to the occurrence of "errors" and "error correction" across five predefined positions within the run, relative to marked deviations from temporal regularity. Efficient production of temporal regularity should engage mechanisms that ensure that the performance in response to an error, that is, a decoupling from the intended rate during a deviation, is adjusted or recoupled with the initial performance. Accordingly, these variables were considered as a more detailed impression of task performance at the local level, spanning a fraction of the entire run. Taken together, global and local measures were considered as representative for the ability to produce a regular sequence of events.

In the perceptual domain, the premise of a sensory experimental setup requires some online measure of covert behavior to assess the efficient perception of temporal regularity. The current study applied ERPs of the EEG, which provide time-sensitive correlates of neurocognitive behavior. Early sensory ERPs (P50 and N100) can be modulated by a sound that deviates from previously established formal regularity (Schwartze, Farrugia, & Kotz, 2013; Bendixen et al., 2012; Slabu, Escera, Grimm, & Costa-Faidella, 2010). Under specific circumstances, the amplitude of these components also reflects an inverse relationship to temporal regularity and predictability, that is, the amplitude is smaller for fully predictable, temporally regular stimuli in contrast to a larger amplitude elicited by temporally irregular stimuli (Schwartze et al., 2013).

Combining the SMT measures and ERPs obtained in temporally regular and irregular contexts, the current study investigates the timing of regular sequences in a sensorimotor and in a purely perceptual task to (i) replicate previous findings in a participant group with a broad age range in order to then (ii) explore potential patterns of covariation between the two domains. On this basis, any systematic pattern of covariation could be interpreted as evidence in favor of common mechanisms underlying the ability to time regular sequences in production and perception, which, in turn, instantiates a precursor for the efficient predictive adaptation to a dynamic environment.

METHODS

Participants

Forty right-handed participants (15 women, mean age = 51.7 years, SD = 12.8 years, range = 26–78 years) with no history of neurological disorder took part in the study. All participants gave their informed written consent and received a compensatory fee. The study was approved by the ethics committee of the University of Leipzig.

Spontaneous Motor Tempo

The SMT of each participant was assessed in two runs, before and after the EEG session. None of the participants had prior experience with this type of task. The first run served as a "target," as it was assumed to reflect the most naturalistic performance obtainable. Data from the target run were later correlated with the EEG data. The second run served as a "control," as it aimed to verify SMT as a stable measure. During both runs, 31 taps were recorded (corresponding to 30 intertap intervals [ITIs]). All participants tapped with the index finger of their right hand on a touch-sensitive rubber surface of an electronic drum pad (SPD-6, Roland Corp., Hamamatsu, Japan), which was connected to the MIDI port of a PC running custom software written in MAX (cycling74.com). Participants were asked to tap for a short while as regularly as possible at a self-paced rate after they had familiarized themselves with the setup and task. A single piano tone was presented via headphones (HD 202, Wedemark, Germany) after 31 taps
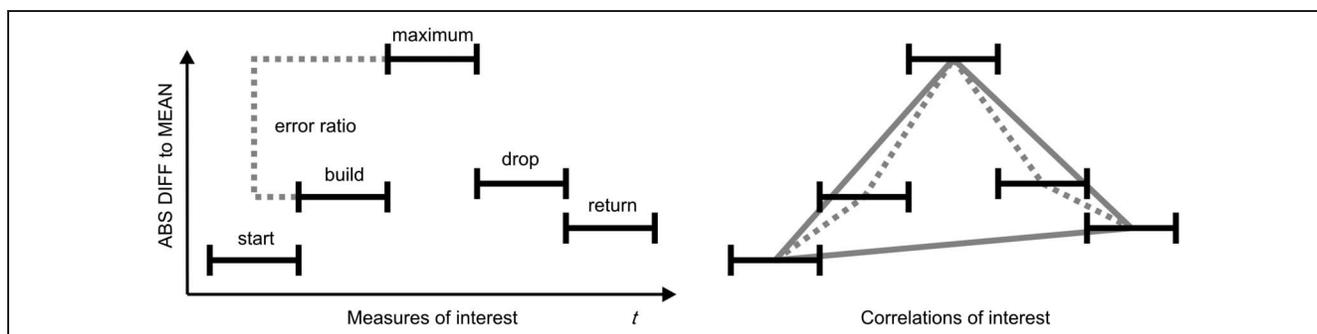
Figure 1. Peak-complex and correlations of interest. The peak-complex (left) provides a number of local measures of task performance during the SMT assessment. It is defined relative to the maximum in terms of the absolute difference (in milliseconds) between a specific ITI and the mean ITI of each run. Two preceding and two following ITIs are considered. Taken together, the five ITIs are interpreted functionally as *start*, *build*, *maximum*, *drop*, and *return*. An additional combined value (*error ratio*, *build/maximum*) is calculated to obtain a relational measure for subsequent use in the correlation analyses with the EEG data. Within the peak-complex (right), two types of correlations of interest are analyzed: linear (dotted lines) and nonlinear (solid lines). The former reflect relations over time, whereas the latter are considered to reflect potential functional relations such as the connection between *start* and *maximum* or between *start* and *return*, that is, the relation between initial performance and maximal deviation and between initial performance and the outcome of error correction.

were registered to signal the end of the run. The use of headphones also served to attenuate any sound that could be generated by the impact of the finger on the pad. There was no training, and no exemplary performance was provided by the experimenter to avoid the introduction of a potential bias.

Before all analyses, missing ITI values corresponding to nonregistered taps because of insufficient force were replaced by values that were obtained via linear interpolation based on the two neighboring ITIs (5.8% of the final data set). Principle variables of interest were the mean (M_ITI) and the coefficient of variation (CV_ITI, *SD* of all ITIs divided by M_ITI) that were calculated separately for each participant and for both runs. The respective values provide a global impression of task performance in terms of rate and variability. Although SMT is considered to reflect self-chosen temporal characteristics, that is, unpaced performance, the instructions required participants to tap as regular as possible. This explicit instruction implied some form of monitoring of performance, which, in turn, suggests that corrective mechanisms should become active in the case of an error. To account for related phenomena, a number of additional variables was analyzed and subsumed under the label "peak-complex" (Figure 1).

The peak-complex is essentially a theory-driven approach to identify a local pattern embedded in the longer SMT run. The rationale behind it is based on several assumptions: (i) although SMT is a form of unpaced spontaneous behavior, it engages error correction in compliance with task instructions; (ii) the largest (sudden or gradual) deviation from the M_ITI is, most likely, related to the occurrence of an error; and (iii) efficient error correction quickly restores pre-error performance, but it is not necessarily instantaneous, that is, it affects the ITI immediately following an error but potentially also subsequent ITIs. In theory, optimal SMT performance would lead to the production of an isochronous sequence. In practice, variable degrees of deviation from such a hypothetical performance

are to be expected. To identify an individual peak-complex, absolute differences between the M_ITI, assumed to reflect the intended tempo as close as possible, and each individual ITI were calculated for each participant. The use of absolute differences served to combine the two principle directions of deviations from the M_ITI (shorter vs. longer ITIs) into one comparable dimension. On this basis, two preceding and two following "functional" ITIs were selected relative, and in addition, to the maximal absolute difference for each participant. Two rather than one ITI were selected to assess processes developing over more than one ITI in both the occurrence of an error and in error correction. From this perspective, *start* most closely approximates the initial pre-error performance, *build* refers to the emergence of an error, *maximum* denotes the maximal deviation, *drop* refers to the reduction of deviation due to corrective mechanisms, and *return* most closely approximates the outcome of error correction. It is important to note that *maximum* is not necessarily equivalent to the actual occurrence of an error, because detected errors in sensorimotor synchronization with a pacing sequence typically lead to overcorrection (Repp & Keller, 2004; Praamstra, Turgeon, Hesse, Wing, & Perryer, 2003; Thaut, Miller, & Schauer, 1998). Accordingly, it is possible that *maximum* already depicts aspects of error correction.

In combination with the global M_ITI and CV_ITI measures, these additional variables were used to obtain an extended set of metrics, which established a relatively detailed individual imprint of the precision of temporal processing during the production of regular temporal structure based on both global and local measures.

EEG

The EEG was recorded with a sampling rate of 500 Hz, an anti-aliasing filter of 135 Hz, and a mastoid online reference from 25 Ag/AgCl scalp electrodes mounted into

an elastic cap according to the standard 10–20 International System. Horizontal and vertical electrooculography was monitored via electrodes placed left and right of the eyes and above and below the right eye, respectively. The ground electrode was placed on the sternum. Electrode impedance was kept below 5 Ω. During the recording, participants sat in a dimly lit, sound-attenuated booth, fixating an asterisk displayed on a screen placed in front of them. The stimulus sequences consisted of 450 sinusoidal tones (standard-to-deviant ratio 4:1, 600-Hz standards, 660-Hz deviants, 300-msec duration, 10-msec rise and fall) presented either at a constant 900-msec tempo (600-msec ISIs) or with ISIs randomly chosen (uniform distribution) from a range between 200 and 1000 msec (Figure 2). The participants were asked to pay attention to the sequence, which was presented via two loudspeakers, to silently count the deviant tones ($n = 90$), and to report the result at the end of the session. A short sequence comprising five deviant tones was attached to the irregular sequence to obtain different correct results for the counting task. Responses to these additional tones were excluded from the EEG analyses. The order of the presentation of the two sequences was counterbalanced across participants. Presentation 12.0 (Neurobehavioral Systems, Albany, CA) was used to control the stimulus delivery. Before the actual recording, a short excerpt from the sequence was used to familiarize the participants with the two tones. Pseudorandomization ensured that the test sequence always started with four instances of a standard tone and that no more than two deviant tones were presented in a row.

EEP 3.2 (Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany; commercially available as EEProbe, ANT Neuro) was used for data preprocessing and analyses. Relatively narrow band-pass filters (about 10–50 Hz) are commonly applied to analyze the P50 component in the context of sensory gating (Chang, Gavin, & Davies, 2012; Patterson et al., 2008). However, for the current combined analysis of P50 and N100, raw data were band-pass filtered from 5 to 75 Hz after re-referencing to averaged mastoids. An automatic rejection algorithm scanned for artifacts exceeding 30 μV at eye channels or 40 μV at electrode CZ. Epochs lasting from −75 to 175 msec relative to stimulus onset were considered for the subsequent analyses. Epochs for standards and de-

viants, which followed the presentation of a deviant tone, were generally rejected. The remaining data were averaged per participant and for the whole group.

The same type of 2 × 2 × 2 × 2 ANOVA with factors Temporal Structure (regular vs. irregular stimulus presentation), Formal Structure (standard vs. deviant), Hemisphere (left vs. right), and Region (anterior vs. posterior) was conducted using SAS 9.3 (SAS Institute, Inc., Cary, NC) for time-windows lasting from 46 to 80 msec (P50) and 80 to 136 msec (N100) relative to stimulus onset in four ROIs. These covered left anterior (F7, F3, FT7, FC3), right anterior (F4, F8, FC4, FT8), as well as left posterior (T7, C3, CP5, P3) and right posterior (C4, T8, P4, CP6) electrode positions to assess the typical frontocentral scalp distribution of the ERPs. In addition, the corresponding CV was calculated for each component and each participant to obtain comparable measures of variability for the SMT run and the ERPs.

RESULTS

Spontaneous Motor Tempo

Global SMT performance (Figure 3) proved similar for the target (M_ITI: 496 msec, $SD = 135$ msec; CV_ITI: .27, $SD = .01$) and the control run (M_ITI: 517 msec, $SD = 151$ msec; CV_ITI: .29, $SD = .03$). The M_ITIs produced in both runs ranged from 209 to 1104 msec. Paired samples t tests comparing the respective measures across the two runs yielded no significant differences for M_ITI ($t(39) = -.87, p = .39$), whereas the same comparison approached significance for CV_ITI ($t(39) = -2.02, p < .06$), indicating a trend toward increased variability during the control run. Significant positive correlations were obtained for M_ITI ($r = .43, p < .01$) and CV_ITI ($r = .31, p < .05$) across the two runs. It is important to note that neither M_ITI (target run: $r = .24, p = .14$; control run: $r = .18, p = .27$) nor CV_ITI (target run: $r = -.03, p = .87$; control run: $r = -.15, p = .36$) was significantly correlated with age. However, taken together, the two variables representing global aspects of the target and the control assessment support the notion of SMT as a relatively stable measure, although the numerically slower rate and the trend toward higher variability in the control run may hint at an

Figure 2. EEG/ERP stimulus sequences. Participants listened to continuous "oddball" sequences consisting of equidurational (300 msec, 10 msec rise and fall times) frequent standard (600 Hz, $n = 360$) and pseudorandomly interspersed infrequent deviant (660 Hz, $n = 90$) sinusoidal tones. The tones were either

separated by a 600-msec ISI (regular condition) or randomly chosen 200–1000 msec ISIs (irregular condition). Four standards were presented at the beginning of the sequence to allow establishing a memory trace.
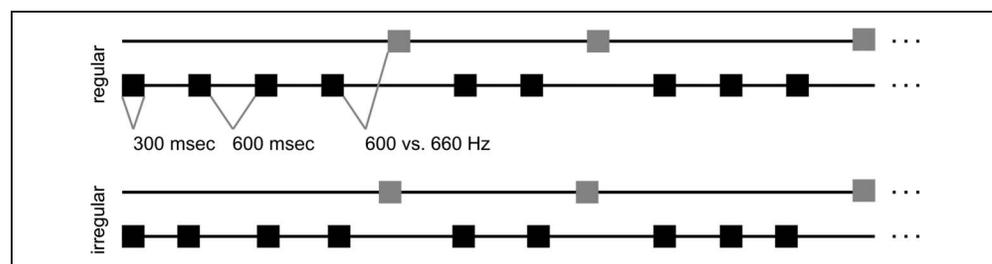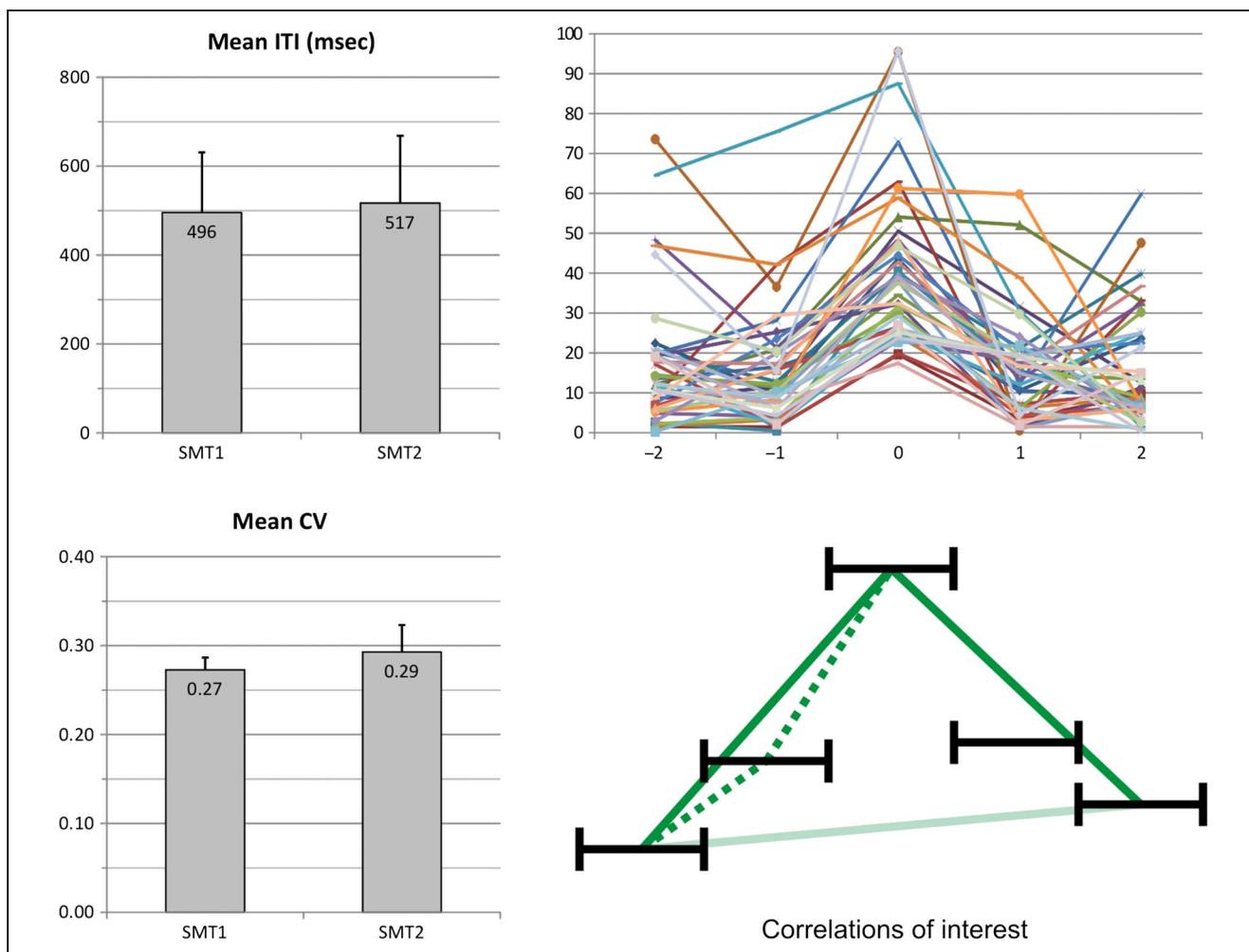
Figure 3. Global and local measures for SMT. The rate (mean ITI) and variability (mean coefficient of variation of ITIs) are comparable for the assessment of SMT recorded before (SMT1, target run) and after (SMT2, control run) the intermediate EEG session (left). For the purpose of visualization, linear regression lines were applied to the data to illustrate the relation of the ITIs, which constitute the peak-complex (top right). Significant relations and a trend toward significance in the correlations of interest are marked in green and light green, respectively (bottom right).

influence of the intermediate EEG session, which employed a considerably slower presentation rate as compared with the M_ITI obtained in both SMT runs.

The maximal absolute difference served to temporally align the primary peak-complex for each participant (Figure 3). If part of the primary complex would have extended beyond the available observations, that is, if the maximum was located in the periphery of the run, the secondary maximum was used and so forth. A potential complex was rejected if it would have overlapped with a rejected peripheral complex or if it would have included more than one interpolated value. In one case, the secondary maximum was used, because the participant had produced two numerically identical neighboring primary maxima.

Visual inspection of the peak-complex for each individual indicated that not all participants followed the expected pattern of an increase in the deviation from *start* over *build* to *maximum*. For the progression from *start* to *build*, the pattern was mixed, with 17 participants

showing the expected increase, as opposed to 16 participants who showed a decrease, and 7 participants who showed comparable performance (<25% difference between *start* and *build*). Although this finding could be seen as weakening the proposed functional interpretation of the corresponding ITIs, it does not contradict it and it is compatible with the notion of sudden as well as more gradual deviations leading to an "error" in SMT. However, the observed variability warrants the additional use of the relative measure of *error ratio* in this context.

The first level of analysis concerned the linear correlations of interest, reflecting the potential direct connections of the respective ITIs over time. This analysis yielded a positive correlation between *start* and *build* ($r = .64$, $p < .001$), as well as between *build* and *maximum* ($r = .68$, $p < .001$). However, no significant results were obtained for the progression from *maximum* to *drop* ($r = .21$, $p = .18$) or the progression from *drop* to *return* ($r = -.11$, $p = .51$). These findings confirm linear

coherence for the initial phases. Larger deviation from the hypothetical standard already at *start* predicts larger deviation at *build*, and in turn, larger deviation at *build* predicts larger deviation at *maximum*. Following maximum, this type of linear coherence is lost, potentially reflecting the actual occurrence of an error and subsequent error correction. Critically, because of the use of absolute differences, the linear coherence identified for the initial stages may reflect the general influence of the M_ITI on the functional ITIs. Indeed, although there was no significant correlation between the two global measures M_ITI and CV_ITI ($r = .21$, $p = .19$), there was a significant correlation between M_ITIs and the corresponding SD ($r = .75$, $p < .001$). However, this type of global influence does not explain why the linear coherence is selectively lost following *maximum*.

The second level of analysis concerned the nonlinear correlations of interest, reflecting the proposed distal functional connections. There were two significant correlations, between *start* and *maximum* ($r = .70$, $p < .001$) and between *maximum* and *return* ($r = .42$, $p < .01$). The correlation between *start* and *return* approached significance ($r = .31$, $p = .052$). Analyses involving the ITI thought to reflect effective reduction of deviation (*drop*) did not yield any significant result. This pattern confirms that larger deviation already at *start* is associated with larger deviation at *maximum*, whereas the trend

toward correlation between *start* and *return* may indicate the effort to return to the initial performance. Decoupled performance on *drop* may reflect interference by corrective mechanisms. The link between slower tapping rate and the larger deviation at *maximum* is probably not trivial, as slower tempi may afford better monitoring of the sequence, which may in turn prevent larger deviations. However, this seems not to be the case.

ERPs

The results obtained for the counting task employed during the EEG recordings confirmed that the participants paid attention to the stimulus sequence in the regular (mean = 89.2, $SD = 1.5$; mean percentage accuracy = 99.1, $SD = 1.7$) and in the irregular condition (mean = 93.1, $SD = 1.9$; mean percentage accuracy = 98.0, $SD = 2.0$). In line with previous findings, initial visual inspection of the EEG data confirmed the expected differentiation of ERP responses reflecting temporally regular and irregular stimulus presentation and the contrast between standard and deviant stimuli, respectively (Figure 4). Accordingly, the ANOVAs conducted for the P50 and N100 time-windows yielded significant main effects of Temporal Structure for the former ($F(1.39) = 22.25$, $p < .0001$) and for the latter component ($F(1.39) = 12.81$,

Figure 4. Group-averaged and individual ERPs. The auditory oddball sequences evoked reliable P50 and N100 ERPs, which display a typical frontocentral scalp distribution and a differentiation of formal structure (standards vs. deviants; ROIs: LA = left anterior, LP = left posterior, RA = right anterior, RP = right posterior). Peak amplitudes for individual responses obtained from single midline electrodes (FZ for P50, CZ for N100) were used in the correlation analyses. For graphical display only, a 30-Hz low-pass filter was applied to the data.
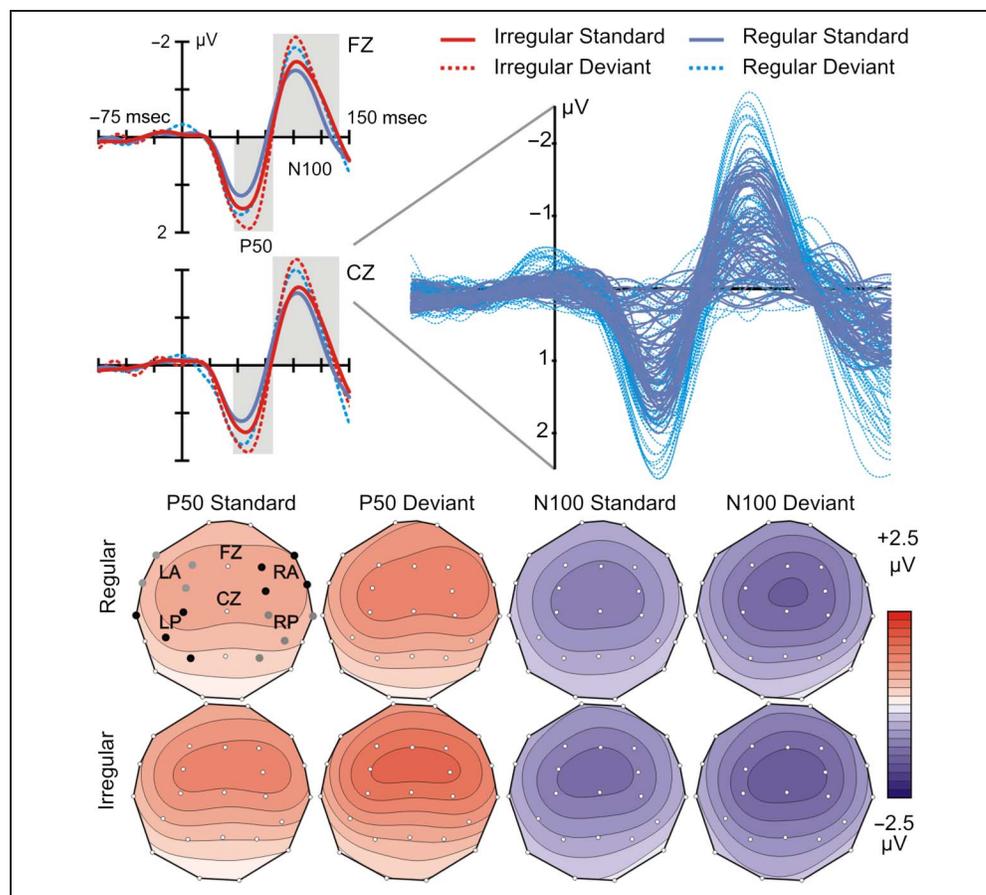
$p < .001$). This was paralleled by significant main effects of Formal Structure for the P50 ($F(1.39) = 32.87, p < .0001$) and the N100 ($F(1.39) = 15.35, p < .001$). In the P50 time-window, there was also an effect of Hemisphere ($F(1.39) = 16.98, p < .001$) and of Region ($F(1.39) = 43.73, p < .0001$), as well as an interaction of Region with each of the other factors, Temporal Structure × Region ($F(1.39) = 8.48, p < .01$), Formal Structure × Region ($F(1.39) = 4.61, p < .04$), Hemisphere × Region ($F(1.39) = 6.24, p < .02$), whereas none of the remaining interactions yielded a significant result (all $ps > .08$). Resolving the significant interactions by Region indicated effects of Temporal Structure in anterior ($F(1.39) = 28.68, p < .0001$) and posterior ($F(1.39) = 13.05, p < .001$) regions, of Formal Structure in anterior ($F(1.39) = 32.27, p < .0001$) and posterior ($F(1.39) = 24.53, p < .0001$) regions, and of Hemisphere in anterior ($F(1.39) = 28.27, p < .0001$) and posterior ($F(1.39) = 5.61, p < .03$) regions.

In the N100 time-window, there was also a significant effect of Region ($F(1.39) = 17.60, p < .001$) and an interaction of Hemisphere × Region ($F(1.39) = 8.09, p < .01$), but no significant effect of Hemisphere ($F(1.39) = 1.29, p = .26$). The interaction Formal Structure × Hemisphere approached significance ($F(1.39) = 3.83, p = .057$), whereas the remaining interactions were nonsignificant (all $ps > .12$). Resolving the significant interaction by Region yielded an effect of Hemisphere in anterior ($F(1.39) = 6.33, p < .02$), but not in posterior regions ($F(1.39) = .18, p = .68$).

Taken together, these findings provide evidence for the differential sensitivity of P50 and N100 amplitudes to temporal and formal structure. More specifically, the results fulfill the secondary goal of the current study and confirm that temporally regular stimulus presentation leads to smaller ERP amplitudes, substantiating the notion of an inverse relation between ERP amplitude and stimulus predictability. Furthermore, there was no significant correlation between age and ERP amplitudes (all $ps > .28$). However, to address the primary goal of the study, that is, the potential relation of production and perception of temporal regularity, subsequent analyses focused exclusively on P50 and N100 measures obtained with regular stimulus presentation and their covariation with the respective SMT measures.

Analogous to the procedure applied to the SMT data, initial correlation analyses tested for linear connections between the P50 and N100 components. Significant negative correlations were obtained for both standards ($r = -.807, p < .001$) and deviants ($r = -.684, p < .001$), indicating that more positive P50 deflections were associated with more negative N100 deflections. It is important to note, however, that with the continuous unfolding of temporal structure inherent to the setup employed during the EEG session, temporal progression does not necessarily conform to a functional relation over time. Rather, it may be the case that modulations of both ERPs reflect "early" stimulus-driven aspects as well as a "late"

influence of predictive adaptation over the course of many stimulus repetitions.

Action–Perception Linking

Although the results from both the sensorimotor SMT and the perceptual EEG/ERP session confirm previous findings in these domains, the critical analyses concerning an individual's capacity for efficient timing across production and perception requires a cohesive pattern of covariation. The first step of the analyses concerned the relation between the ERP amplitude measures and the global SMT variables. All ERP measures correlated significantly with M_ITI (Figure 5). More specifically, significant positive correlations were obtained for the P50 for standards ($r = .51, p < .01$) and the P50 for deviants ($r = .41, p < .01$), whereas negative correlations were obtained for the N100 for standards ($r = -.40, p < .02$) and the N100 for deviants ($r = -.44, p < .01$). This pattern confirms larger ERP deflections, that is, more positive P50 amplitudes and more negative N100 amplitudes, with slower tapping rates. However, none of the ERP amplitude measures was significantly correlated with CV_ITI, the global index of efficient task performance in terms of temporal regularity (all $ps > .52$), although there were indications for covariation across the global CV measures for the P50 for standards ($r = .32, p < .05$), but not for the remaining components (all $p > .11$), confirming that in this case more variable production was associated with more variable responses to the standard tones.

The second step of analyses concerned the relation between *error ratio* and the ERP measures. This local measure of task performance correlated significantly with the P50 for deviants ($r = .35, p < .03$), the N100 for standards ($r = -.33, p < .04$), and the N100 for deviants ($r = -.43, p < .01$), whereas there was potentially a marginal trend in the case of the P50 for standards ($r = .27, p = .096$). Nevertheless, for the former ERPs these results confirm larger amplitudes with a proportionally larger relative deviation from the intended performance, that is, less efficient local performance.

The final analyses explored the relation between the ERP measures and the ITIs of the peak-complex. All ERP measures correlated with *start*: P50 for standards ($r = .35, p < .03$), P50 for deviants ($r = .41, p < .02$), N100 for standards ($r = -.33, p < .05$), N100 for deviants ($r = -.36, p < .03$). Similar results were obtained for *build*: P50 for standards ($r = .35, p < .03$), P50 for deviants ($r = .40, p < .02$), N100 for standards ($r = -.33, p < .04$), N100 for deviants ($r = -.41, p < .01$). However, none of the subsequent intervals correlated with the ERP measures, that is, the correlations with *maximum*, *drop*, or *return* did not yield any significant result (all $ps > .10$). Although this pattern of results confirms action–perception linkage during pre-error performance, this initial coherence is lost with the occurrence of an

Figure 5. Correlations across production and perception of regular temporal sequences. (A) P50 and N100 amplitudes are significantly correlated with M_ITI. (B) Removal of outliers in this data set, defined as the two fastest and the two slowest performances did not substantially change the pattern. (C) Although the global CV_ITI was not significantly correlated with any ERP measure, the local measure of *error ratio* was, with the exception of the P50 for standards. (D) Results for the peak-complex demonstrate significant correlations between SMT performance and ERP amplitude during *start* and *build*, but not during the following ITIs, potentially because of the engagement of additional/different mechanisms responsible for error correction during these stages. *ns* = not significant.

error, potentially as a result of the recruitment of corrective mechanisms.

DISCUSSION

The primary goal of the current study was to investigate the human ability to generate a temporally regular sequence of events in relation to the ability to perceive a temporally regular sequence of events. The secondary goal was to further explore the inverse relation of early auditory ERPs and temporal predictability in a population with a broad age range. These goals were assessed on the basis of global and local characteristics of repetitive spontaneous finger-tapping and evoked responses to sinusoidal tones presented in temporally regular and irregular contexts. The findings obtained in each domain are in line with previous results (Schwartze et al., 2013; Schwartze, Keller, Patel, & Kotz, 2011), whereas the linking of both domains revealed additional cross-domain covariation. These latter findings seem compatible with the notion of a more general capacity for efficient temporal processing.

Overall, SMT proved to be a stable measure, although the performance during the second run was probably not independent of the global temporal context of the

experimental setup (McAuley & Miller, 2007). A bias that was introduced by the exposure to the slower stimulus presentation rate employed during the EEG session could explain the numerically lower rate and the trend toward higher variability that was observed for the second SMT run. The current results confirm previous reports of relatively high interindividual variability as opposed to low intraindividual variability (Fraisse, 1982). The group-averaged M_ITI rates were somewhat faster than the classical value proposed (600 msec), with an average M_ITI across both runs centered at 507 msec. This observation is in line with our own work testing a similar age range as well as a comprehensive evaluation of preferred tempo, which suggests a value around 500 msec (Schwartze, Keller, et al., 2011; Moelants, 2002; but see Baudouin, Vanneste, & Isingrini, 2004; Drake, Jones, & Baruch, 2000). However, the current findings also demonstrate that, despite its apparent simplicity, SMT is a multilayered phenomenon.

Although the theoretically motivated peak-complex is but an approximation of some of its facets, it seems to be a valid construct, as it yields a number of critical findings that are not captured by global measures such as M_ITI and CV_ITI. Overall, these empirical findings support most of the original theoretical assumptions. The occurrence of a deviation from the hypothetical intended

rate seems to interrupt coherence of performance over time, which is then almost restored within two ITIs. However, the current setup does not permit speculation concerning the nature of the interruption. Error correction is a likely candidate, but as opposed to paced tapping performance, it is unclear to what extent error correction in SMT would engage, for example, motor control, attention, intention, or awareness of the deviation (Repp & Keller, 2004). Moreover, by definition, the peak-complex reflects a local pattern. A more realistic account would have to consider different profiles and different types of both errors and error correction, for example, modeled after the distinction of phase shifts, step changes, and event onset shifts in sensorimotor synchronization (Repp, 2002). Ultimately, this would require classification of patterns extracted from substantially longer time series. This is, however, outside the scope of the current study.

Like the SMT recordings, EEG results confirmed previous reports concerning the effects of temporal and formal predictability within the relatively early time range covered by the P50 and the N100 components. In this regard, it is important to acknowledge that these ERP amplitudes alone are a highly selective marker of perceptual efficiency. Although their inverse relation to temporal predictability offers a starting point to explore the use of temporal regularity in perception, it is obvious that the quality of predictive adaptation as a whole depends on the interplay of numerous processes, ranging from peripheral to higher-level controlled processes and vice versa.

Individual tapping rates, expressed in terms of M_ITI, were significantly correlated with ERP amplitudes, confirming a basic pattern of similarity between motor and perceptual electrophysiological measures. This was not the case for CV_ITI, the global measure of variability and efficient task performance, although there was selective covariation between variability in the P50 in response to standards and CV_ITI, indicating that more variable ERP amplitudes were paralleled by higher variability in the production of temporal regularity. Considering that responses to standards may reflect temporal regularity most clearly, that is, irrespective of formal stimulus aspects, this finding seems particularly relevant. However, the local measure of task performance, *error ratio*, was significantly correlated with all ERPs, except for the P50 in response to standard tones, which showed a nonsignificant trend in the same direction. The overall picture suggests that these results are not primarily driven by "deviance," but conform to standards and deviants, strengthening the argument for a basic function of temporal regularity in perception. Nevertheless and more specifically, the question arises as to how different aspects of deviance processing in terms of SMT compare to different hierarchical and functional stages of deviance processing and distraction (Grimm & Escera, 2012; Horvath, Winkler, & Bendixen, 2008). A better understanding of the underlying mechanisms and the fact that SMT recordings are independent

of external stimulation may also establish an additional link toward predictions about the perceptual consequences of self-generated actions and internal forward modeling, functions that are associated with the cerebellum, one of the essential nodes of the proposed dedicated temporal processing system (Knolle, Schröger, & Kotz, 2013; Wolpert, Miall, & Kawato, 1998). In this context, it may be critical to consider the time course determined by the circular nature of repetitive sensorimotor behavior and to dissociate between first-pass processes related to, for example, the first initiation of action, and later processes, which may reflect the influence of predictive adaptation or the updating of an internal model.

The results obtained from the combined analysis of ERPs and the peak-complex provide further evidence for a cohesive pattern of covariation as well as a hint toward a dissociation of sensory and motor aspects of SMT. The strong trend in the correlation of *start* and *return* in the separate SMT analysis suggests that, following a deviation, initial performance can be quickly restored. However, the correlation of the ERP measures with the performance at *start* and *build* is lost with maximal deviation and not reestablished at *return*. This pattern may indicate sustained engagement of controlled mechanisms to monitor production following an error and the actual error correction.

Marked pathological changes to the temporal processing system lead to impaired production and perception of temporal structure and potentially to general problems in adapting to a changing environment, for example, in Parkinson disease, Huntington disease, or following BG lesions (Cope, Grube, Singh, Burn, & Griffith, 2014; Allman & Meck, 2012; Schwartze, Keller, et al., 2011). As inefficient temporal processing and use of temporal regularity may contribute to a broad range of cognitive behaviors, this link may also be reflected in the relation of tapping ability and several higher-level functions, including linguistic skills (Tierney & Kraus, 2013a, 2013b). In turn, such coupling may offer a means to optimize cognitive processes by manipulating the temporal structure and regularity of information, potentially also via the training of sensorimotor skills and the transfer to seemingly unrelated tasks and domains. However, about 4% of the general population demonstrate selective difficulties in synchronizing movements to more complex musical sequences, whereas synchronization to simple metronome sequences is unimpaired (Sowiński & Dalla Bella, 2013). Although these difficulties may be a consequence of inefficient temporal processing, it stands to reason as to what extent, how, and under what circumstances low-level processes such as the precise timing of temporal regular sequences may influence higher-level functions such as music and speech processing. Irrespective of these open questions, it is evident that events do not simply unfold in time. Rather, temporal structure and the respective temporal information should be considered relevant by their own virtue. Accordingly, the capacity for efficient

temporal processing and timing of regular sequences may represent a common denominator across different domains, serve predictive adaptation, and express itself in the relation of simple experimental measures such as SMT and auditory ERPs.

REFERENCES

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain, 135,* 656–677.

Baudouin, A., Vanneste, S., & Isingrini, M. (2004). Age-related cognitive slowing: The role of spontaneous tempo and processing speed. *Experimental Aging Research, 30,* 225–239.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology, 83,* 120–131.

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience, 20,* 204–214.

Chang, W., Gavin, W. J., & Davies, P. L. (2012). Bandpass filter settings differentially affect measurement of P50 sensory gating in children and adults. *Clinical Neurophysiology, 123,* 2264–2272.

Cope, T. E., Grube, M., Singh, B., Burn, D. J., & Griffith, T. D. (2014). The basal ganglia in perceptual timing: Timing performance in Multiple System Atrophy and Huntington's disease. *Neuropsychologia, 52,* 73–81.

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "what" and "when" in the auditory system: Temporal predictability enhances repetition suppression. *The Journal of Neuroscience, 31,* 18590–18597.

Coull, J. F., Cheng, R., & Meck, W. H. (2010). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology, 36,* 3–25.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition, 77,* 251–288.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.

Fuster, J. M. (2001). The prefrontal cortex—An update: Time is of the essence. *Neuron, 30,* 319–333.

Fuster, J. M., Bodner, M., & Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature, 405,* 347–351.

Grimm, S., & Escera, C. (2012). Auditory deviance detection revisited: Evidence for a hierarchical novelty system. *International Journal of Psychophysiology, 85,* 88–92.

Horvath, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology, 79,* 139–147.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 3–18.

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex, 49,* 2449–2461.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology, 13,* 250–255.

MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology, 99,* 1164–1173.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General, 135,* 348–367.

McAuley, J. D., & Miller, N. S. (2007). Picking up the pace: Effects of global temporal context on sensitivity to the tempo of auditory context. *Perception & Psychophysics, 69,* 709–718.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36,* 313–336.

Merchant, H., Zarco, W., & Prado, L. (2008). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology, 99,* 939–994.

Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition* (pp. 580–583). Adelaide, Australia: Causal Productions.

Nobre, A. C., Correa, A., & Coull, J. T. (2007). The hazards of time. *Current Opinion in Neurobiology, 17,* 465–470.

Patterson, J. V., Hetrick, W. P., Boutros, N. N., Jin, Y., Sandman, C., Stern, H., et al. (2008). P50 sensory gating ratios in schizophrenics and controls: A review and data analysis. *Psychiatry Research, 158,* 226–247.

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.

Repp, B. H. (2002). Automaticity and voluntary control of phase correction following event onset shifts in sensorimotor synchronization. *Journal of Experimental Psychology, 28,* 410–430.

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.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11,* 211–218.

Schütz-Bosbach, S., & Prinz, W. (2007). Prospective coding in event representation. *Cognitive Processing, 8,* 93–102.

Schwartze, M., Farrugia, N., & Kotz, S. A. (2013). Dissociation of formal and temporal predictability in early auditory evoked potentials. *Neuropsychologia, 51,* 320–325.

Schwartze, M., Keller, P. E., Patel, A. D., & Kotz, S. A. (2011). The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the

detection of tempo changes. *Behavioural Brain Research, 216,* 685–691.

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology, 87,* 146–151.

Slabu, L., Escera, C., Grimm, S., & Costa-Faidella, J. (2010). Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials. *European Journal of Neuroscience, 32,* 859–865.

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51,* 1952–1963.

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.

Tierney, A., & Kraus, N. (2013a). The ability to move to a beat is linked to the consistency of neural responses to sound. *The Journal of Neuroscience, 33,* 14981–14988.

Tierney, A., & Kraus, N. (2013b). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and Language, 124,* 225–231.

Vanneste, S., Pouthas, V., & Wearden, J. H. (2001). Temporal control of rhythmic performance: A comparison between young and old adults. *Experimental Aging Research, 27,* 83–102.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *Neuroimage, 49,* 1728–1740.

Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14,* 5–12.

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences, 13,* 532–540.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2,* 338–347.