CrossMark

BRIEF REPORT

New insights into statistical learning and chunk learning in implicit sequence acquisition

Yue Du[1,2] · Jane E. Clark[1,3]

Abstract Implicit sequence learning is ubiquitous in our daily life. However, it is unclear whether the initial acquisition of sequences results from learning to chunk items (i.e., chunk learning) or learning the underlying statistical regularities (i.e., statistical learning). By grouping responses with or without a distinct chunk or statistical structure into segments and comparing these responses, previous studies have demonstrated both chunk and statistical learning. However, few studies have considered the response sequence as a whole and examined the temporal dependency of the entire sequence, where the temporal dependencies could disclose the internal representations of chunk and statistical learning. Participants performed a serial reaction time (SRT) task under different stimulus interval conditions. We found that sequence learning reflected by reaction time (RT) rather than motor improvements represented by movement time (MT). The temporal dependency of RT and MT revealed that both RT and MT displayed recursive patterns caused by biomechanical effects of response locations and foot transitions. Chunking was noticeable only in the presence of the recurring RT or MT but vanished after the recursive component was removed, implying that chunk formation may result from biomechanical constraints rather than learning itself. In addition, we observed notable first-order autocorrelations in RT. This trial-to-trial association enhanced as learning progressed regardless of stimulus intervals, reflecting the internal cognitive representation of the first-order stimulus contingencies. Our results suggest that initial acquisition of implicit sequences may arise from first-order statistical learning rather than chunk learning.

✉ Yue Du
  duyue@umd.edu

1   Department of Kinesiology, University of Maryland, 2351 School of Public Health Bldg, 4200 Valley Drive, College Park, Maryland 20742-2611, USA

2   Applied Mathematics & Statistics, and Scientific Computation Program, University of Maryland, 3103 Mathematics Building, College Park, Maryland 20742-4015, USA

3   Neuroscience and Cognitive Science Program, University of Maryland, 0218 Cole Student Activities Building, 4090 Union Drive, College Park, Maryland 20742, USA

Introduction

The ability to acquire motor sequences is crucial to our daily activities, such as riding a bicycle, dressing, and driving a car. These sequences are thought to be learned implicitly where an individual does not have the explicit knowledge of the sequence before learning. However, it remains unclear what mechanism underlies the implicit learning of sequences. On one hand, chunk learning has been widely considered to drive the implicit acquisition of cognitive and motor skills (Gobet et al., 2001; Gobet, Lloyd-Kelly, & Lane, 2016). Not surprisingly, implicit learning of motor sequences can result from chunk learning (Boyd et al., 2009; Cohen, Ivry, & Keele, 1990; Koch & Hoffmann, 2000) where a sequence is partitioned into short segments to be learned and the concatenations of segments leads to the acquisition of the sequence.

On the other hand, implicit sequence learning can arise from statistical learning (Hunt & Aslin, 2001)—a process where the probabilistic structure underlying the entire sequence is learned and encoded in the form of an internal model that represents the stimulus contingencies (Bornstein & Daw, 2012; Visser, Raijmakers, & Molenaar, 2007).

To date, chunk and statistical learning have been supported by studies that used the serial reaction time (SRT) task (Nissen & Bullemer, 1987). For example, studies that attempted to demonstrate chunk learning designed a sequence with salient cues that allowed chunk to be formed (Kirsch, Sebald, & Hoffmann, 2010; Koch & Hoffmann, 2000). Similarly, statistical learning has been demonstrated when sequences had definite statistical structures (Hunt & Aslin, 2001). However, it remains unknown whether chunk or statistical learning or both are used to acquire a sequence, especially when the sequence does not have salient cues of chunks and statistical properties.

To answer this question, we employed the SRT task and addressed limitations in previous studies that may confound the findings of chunk and statistical learning. First, in previous studies, chunk or statistical learning was measured by response time. We examined two phases of response time: reaction time (RT), which reflects mental processing (Sternberg, 1969), and movement time (MT), which characterizes the movement itself (Moisello et al., 2009). It is essential to understand whether chunk or statistical learning found in previous studies was attributed to RT and/or MT patterns.

Second, in earlier studies, sequence elements with a chunk or statistical structure were separated from the entire sequence and these separated elements were used to infer chunk or statistical learning. However, this is artificial since the sequence is performed as a whole. Segregating the sequence ruins the temporal dependency of the sequence and thus may mislead our understandings of chunk and statistical learning. In addition, the effects of biomechanical constraints have been ignored when studying chunk or statistical learning. For example, a component of RT and MT may depend on response effectors (i.e., fingers or hands) due to biomechanical constraints of the body. This preexisting and fixed component of RT/MT does not change as sequence learning progresses and it repeats when a fixed sequence recurs. Thus, it is critical to take this fixed RT/MT component into consideration when studying chunk or statistical learning. To circumvent these problems, we examined RT and MT by time series decompositions and fitting them with seasonal autoregressive models. We identified the periodic component of RT/MT and examined whether chunk learning depends on the presence of recursive patterns in RT and MT that carries the information of biomechanical effects. Furthermore, the autoregressions of RT/MT that are separated from the periodic component (exempt from the biomechanical effects) may reflect an internal cognitive representation of the sequence structure (i.e., RT) or trial-by-trial association between movements (i.e., MT).

In addition, we asked participants to perform the SRT task under different stimulus interval conditions. The stimulus interval may be an important factor in chunk learning. For example, the performer's short-term memory that links successive sequence elements could be impaired when these elements are separated by a long time interval (Frensch & Miner, 1994). Likewise, evidence from visual and speech learning studies has suggested that statistical learning also may be subjected to the stimulus interval effect (Toro, Sinnett, & Soto-Faraco, 2005; Turk-Browne, Junge, & Scholl, 2005). Thus, it raises the possibility that chunk learning and statistical learning may vary depending on the stimulus interval conditions.

Methods

Participants

Thirty, nonmusician adults (age: 20.4 ± 0.29 years, 18 females) without neurological disorders signed consent form and participated in this study. The study was performed in accordance with the guideline approved by the Institutional Review Board at University of Maryland, College Park.

Methods and procedure

Participants performed a modified serial reaction time (SRT) task. They were instructed to step to a spatially matched target as quickly as they could when one of six visual stimuli appeared on the computer screen and then step back to the home position (Fig. 1a). The SRT task was performed under one of three stimulus interval conditions (10 participants were randomly assigned to each condition). In condition I, each stimulus was presented for 700 ms before its disappearance and the next stimulus appeared after an interval of 600 ms (700 + 600 ms), yielding a 1300-ms-long interstimulus-interval (ISI). The time intervals were 700 + 200 ms for condition II and 300 + 600 ms for condition III; both generated a 900-ms-long ISI. We chose these three time combinations to control the effect of stimulus appearance or disappearance time that may contaminate the effect of total ISI (Supplementary Methods).

After completing a practice block where the stimuli appeared in a random order, participants performed eight learning blocks for their assigned ISI condition. Specifically, in blocks 1-4, 6, and 8, visual stimuli followed 10 repetitions of sequence A (142315246536). There was no inversion (i.e., 123321) or repetition (i.e., 123123) that provides salient cues of chunk boundaries (Jimenez, 2008). In addition, each element appeared equal times and each was followed by two other elements with equal likelihoods. Sequence B (146252341356) was repeated 10 times in block 5. In block 7, we used sequence A while replacing two 12-item trials at the middle and end of this block with sequence
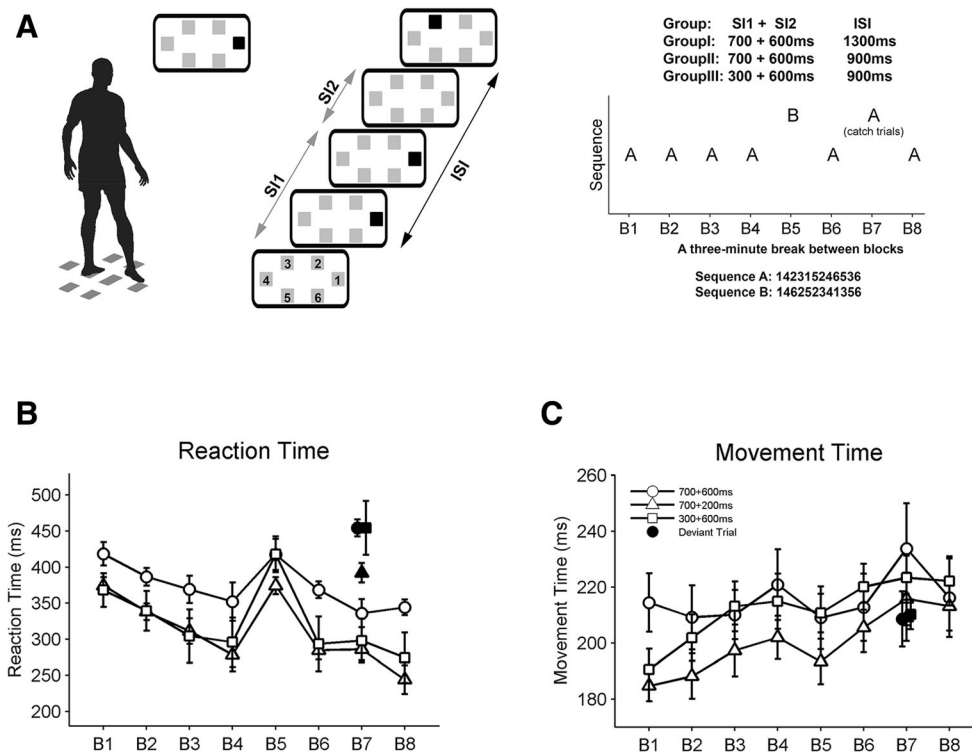
Fig. 1 **a** Experiment procedure. **b** Mean RT across learning blocks. **c** Mean MT across learning blocks. *Error bars* represent standard errors

B as deviant trials. Throughout the task, participants were not instructed about the sequence presentation. A 3-minute rest was provided after each block.

Data analysis

Reaction time (RT) and movement time (MT) were used to measure performance in the task (Supplementary Methods). RT was computed as the time interval between the onsets of visual stimulus and foot movement. MT was quantified as the time discrepancy between the onset and the end point of foot movement. The mean RT and MT were computed for each block. Mean RT/MT differences between blocks 4 and 5 indicated the learning of sequence A (Robertson, 2007). Mean RT/MT differences between blocks 1 and 8 were used as a marker of learning through the entire task. In addition to the learning quantified by performance differences between blocks, we compared performance between sequence A and deviant trials within learning block 7. Specifically, the mean RT/MT on sequence A was computed on 12-step learning trials that preceded the deviant trials.

The sample autocorrelation (ACF) and partial autocorrelation functions (PACF) were used to examine the temporal self-dependency of RT and MT time series. It was clearly shown that a fixed component of RT and MT exhibited periodicity every 12 steps that were identical to the length of sequences A and B. Importantly, this periodic component was confirmed to originate from biomechanical constraints (e.g., the response location

and foot-transition) and it did not affect sequence learning (Supplementary Methods and Results). We then investigated whether chunks were formed with and without the presence of the periodic component. Briefly, we used k-means clustering to group fast and slow responses (Song & Cohen, 2014). Chunk formation was then indicated by the differences in RT between fast and slow responses within each block (Jimenez, 2008). The Pearson's correlation on chunks between learning blocks were computed to quantify the similarity of chunk formation as learning progressed (Song & Cohen, 2014).

To quantify the temporal self-dependency of RT/MT, each time series was then fitted with seasonal autoregressive models. The periodic component was included in the model to account for repeated RT/MT patterns resulting from biomechanical constraints of foot stepping. Thus, the fitted autoregressive coefficients are immune from those biomechanical effects. Fitted autoregressive coefficients were subsequently used for statistical analysis.

Statistical analysis

Two-way mixed-effect ANOVAs were used to examine the effects of learning block, ISI group, and their interaction on mean RT/MT and autoregressive coefficients. A two-way mixed-effect ANOVA was performed to examine the RT/MT difference between learning and deviant trials within block 7. The Dunnett's tests were conducted following a significant effect of ISI group (i.e., the 1300-ms group was the control group). Bonferroni corrected *post hoc* tests were used when

the effect of learning block or trial type is significant. The covariance matrix structure in mixed-effect ANOVAs was determined by the Akaike information criterion (AIC). Student t-tests were used to examine whether the differences between fast and slow RT/MT (i.e., chunk formation) were significantly different from zero for each block and ISI condition (the family-wise error rate was controlled by the Bonferroni correction). The significance level for statistical analyses was set at $\alpha=0.05$.

Results

RTs were significantly affected by block (F(7, 188) = 30.99, $p < 0.0001$, $\eta^2=0.51$) but not by ISI (F(2, 188) = 1.11, $p = 0.33$, $\eta^2=0.0006$) and their interaction (F(14, 188) = 0.84, $p = 0.63$, $\eta^2=0.05$). RT improved from block 1 to other blocks when sequence A was performed (all $p <0.005$) and remained the same between blocks 1 and 5 ($p = 0.92$). RT was faster in block 4 compared with block 5 ($p < 0.0001$). Within block 7, RT was faster in learning than deviant trials (F(1, 27) = 46.12, $p < 0.0001$, $\eta^2=0.52$; Fig. 1b). RT was not affected by ISI (F(2, 27) = 0.35, $p = 0.7$, $\eta^2=0.009$) and its interaction with trial type (F(2, 27) = 1.16, $p = 0.33$, $\eta^2=0.06$). Most importantly, RT, as a part of response time, exhibited the same pattern as response time that improved as learning progressed from block 1 to 4 and became longer in block 5 (Supplementary Results).

Regarding MT, there was a significant effect of block (F(7, 188) = 20.73, $p < 0.0001$, $\eta^2=0.41$) and no effects of ISI (F(2, 188) = 1.16, $p = 0.32$, $\eta^2=0.0003$) and the interaction between ISI group and block (F(14, 188) = 1.4, $p = 0.16$, $\eta^2=0.085$). Unlike RT, MT became longer as learning progressed (Fig. 1c), revealed by the faster MT in block 1 than blocks 3, 4, and 6 to 8 (all $p < 0.05$). MT was comparable between blocks 4 and 5 ($p = 0.2$). Within block 7, MT of deviant trials was faster than that of learning trials (by a magnitude of 16.72 ms; F(1, 27) = 24.25, $p < 0.0001$, $\eta^2=0.44$). MT did not depend on ISI (F(2, 27) = 0.39, $p = 0.68$, $\eta^2=0.003$) and its interaction with trial type (F(2, 27) = 2.58, $p = 0.1$, $\eta^2=0.14$). These results demonstrate sequence learning rather than motor improvements and suggest that the sequence learning is comparable among all ISI groups. The results are consistent with previous studies (Destrebecqz & Cleeremans, 2003; Willingham, Greenberg, & Thomas, 1997).

We examined the sample autocorrelation (ACF) and partial autocorrelation functions (PACF) of RT and MT time series. The ACF tailed off in the periodic lags at 12, 24, 36, and so on (Fig. 2a and b), whereas the PACF cut off after lag 12 (Fig. 2c and d), demonstrating that the fixed RT or MT pattern recurred every 12 steps that was identical to the length of sequences A and B. Importantly, the periodic RT and MT exhibited chunk patterns (Fig. 3a and b) demonstrated in previous studies where a slower response was followed by faster responses

(Koch & Hoffmann, 2000). It was found that chunks were formed in all learning blocks, as well as blocks 5 and 7, as revealed by RT differences between fast and slow responses that were significantly larger than zero (all $p < 0.0001$, all Cohen's d between 1.18 and 2.35; Fig. 3c). The mean Pearson's correlation on 12 mean RTs across learning blocks indicated that chunks started to be formed from block 1 and the same chunks were formed in all learning blocks ($\rho = 0.87$ (mean) ± 0.13 (standard deviation) for 700 + 600 ms ISI; $\rho = 0.77 \pm 0.18$ for 700 + 200 ms ISI; $\rho = 0.77 \pm 0.15$ for 300 + 600 ms ISI). However, after the periodic RT component was removed, the RT differences approached zero (all $p > 0.2$, all Cohen's d between 0.08 and 0.49), indicating the absence of chunk formations. Like RT, chunks were found in MT (all $p < 0.0001$, all Cohen's d between 1.3 and 3.64; Fig. 3d). However, chunks were no longer observed after the periodic MT component was removed (all $p > 0.1$, all Cohen's d between 0.03 and 0.58). These results indicate that chunking was noticeable only in the presence of the recurring RT or MT and thus suggest that chunk formation is likely to result from the fixed periodic component of RT/MT that represents the information of biomechanical effects (Supplementary Results).

In addition to the periodic pattern, the PACF of RT cut off at lag 1, while the ACF gradually tailed off (Fig. 2a and c), representing a significant first-order autoregressive process. However, examinations of individuals' RT revealed that there were individual differences in the order of autoregressions (i.e., order of 1 and 2). This autoregression together with the periodic component suggests that a seasonal ARIMA(p, 0, 0)x(1, 0, 0)$_{12}$ model with p = 1 or 2 is appropriate to describe the RT data. Thus, the AIC was used to select the favored model (between models with a p of order 1 or 2) for each RT time series. We referred to the models as AR(1) or AR(2) model. Unlike RT, MT did not show significant autocorrelations at lags that were less than order 12 (Fig. 2b and d), implying no self-dependencies (besides the periodic pattern) within MT. Thus, we did not model MT for further analyses.

Figure 4a shows the first-order autoregression coefficients that were averaged across individuals' coefficients estimated by their favored model. The coefficient magnitude was significantly affected by block (F(7, 188) = 8.75, $p < 0.0001$ $\eta^2=0.21$) and ISI (F(2, 188) = 9.02, $p < 0.001$, $\eta^2=0.01$) but not their interaction (F(14, 188) = 1.17, $p = 0.3$, $\eta^2=0.01$). The coefficient was smaller when the ISI was 700 + 600 ms than 700 + 200 ms ($p = 0.05$) and 300 + 600 ms ($p < 0.001$). Additionally, all groups increased the coefficient magnitudes from block 1 to 8 ($p < 0.0001$), but there was no difference between blocks 4 and 5 ($p = 0.16$). Student's t tests revealed that the coefficients were significantly larger than 0 in all blocks (all $p < 0.0001$, all Cohen's d between 1.18 and 2.71). These results suggest that although the strength of first-order autoregression between RT depended on stimulus
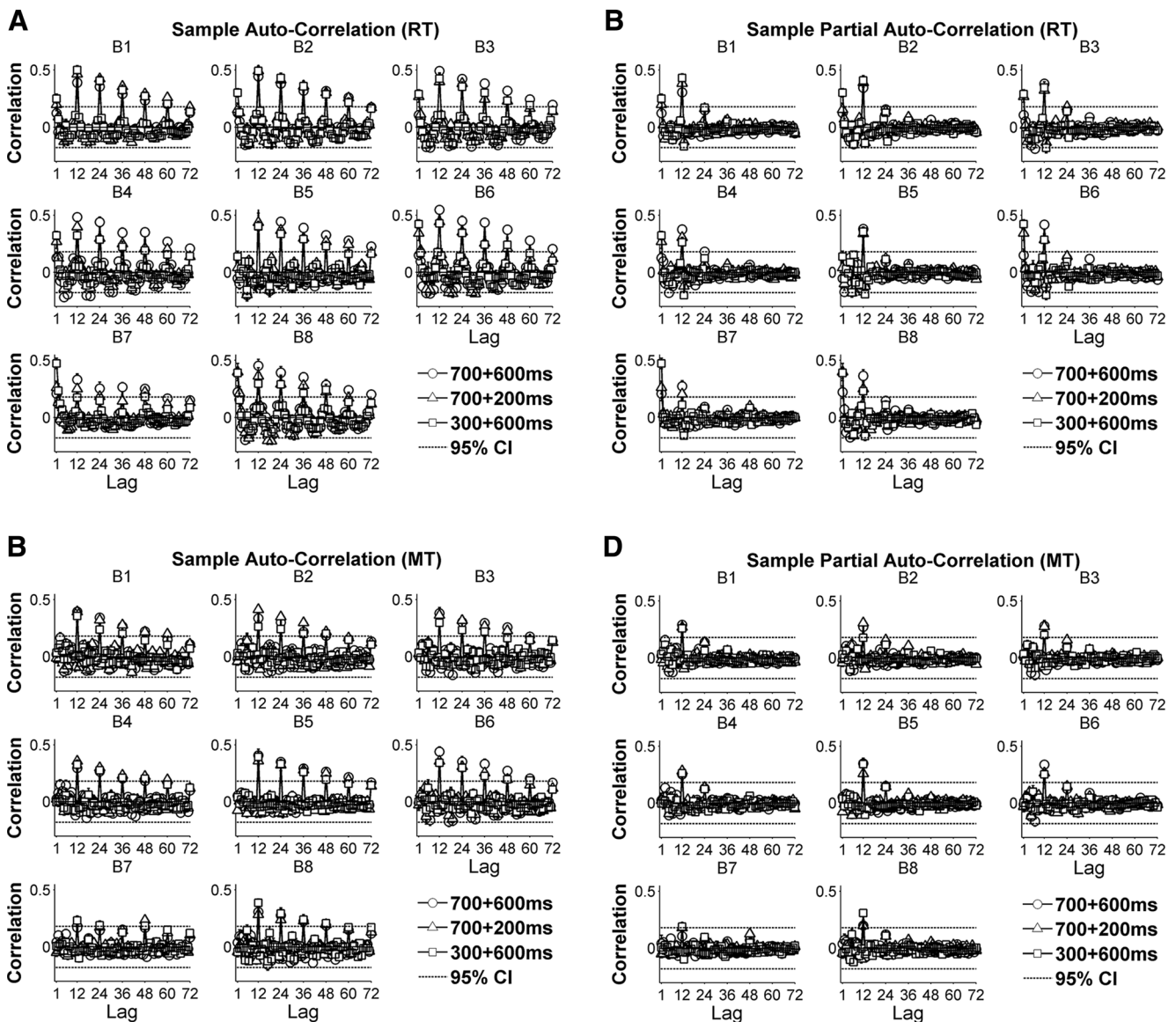
Fig. 2 **a** Mean sample autocorrelation of RT. **b** Mean sample autocorrelation of MT. **c** Mean sample partial autocorrelation of RT. **d** Mean sample partial autocorrelation of MT

intervals, the first-order autoregression strengthened as learning progressed despite the stimulus interval.

Given that only a few participants favored AR(2) model (Fig. 4b), it was impossible to compare the second-order coefficients using individuals' favored models. Thus, we further fitted all RT time series using AR (2) models. The results of first-order coefficients from the AR(2) model did not qualitatively change from coefficients estimated by individuals' favored models that were shown in Fig. 4a. However, there were no effects of ISI ($F(2, 188) = 0.22$, $p = 0.8$, $\eta^2 = 0.0009$), block ($F(7, 188) = 1.82$, $p = 0.09$, $\eta^2 = 0.06$), and their interaction ($F(14, 188) = 0.55$, $p = 0.9$, $\eta^2 = 0.04$) on the second-order coefficients (Fig. 4b). Unlike the first-order coefficients, the second-order coefficients were not significantly different from zero in all blocks (all $p > 0.1$, all Cohen's d between 0.09 and 0.5). These results taken together suggest that implicit

sequence learning may arise through first-order statistical learning.

Discussion

Using an SRT task with higher movement demands than the typical finger SRT tasks, we demonstrated that the early acquisition of an implicit sequence is best characterized as first-order statistical learning rather than the learning of movement chunks. Specifically, our RT results exhibit trial-to-trial associations that may reflect the cognitive representation of the internal model of stimulus associations. In addition, the chunk patterns in both RT and MT appear to result from physical properties of the body (i.e., biomechanically constrained response tendencies to various response locations or foot-

Fig. 3 a Examples of the periodic components of RT in each group. b Examples of the periodic components of MT in each group. As the sequence repeats, the same pattern of RT or MT reoccurs. For display, only first 60 steps in block 1 were illustrated. c Chunk formation in RT when the periodic component was included and excluded. d Chunk formation in MT when the periodic component was included and excluded. *Error bars* represent standard errors
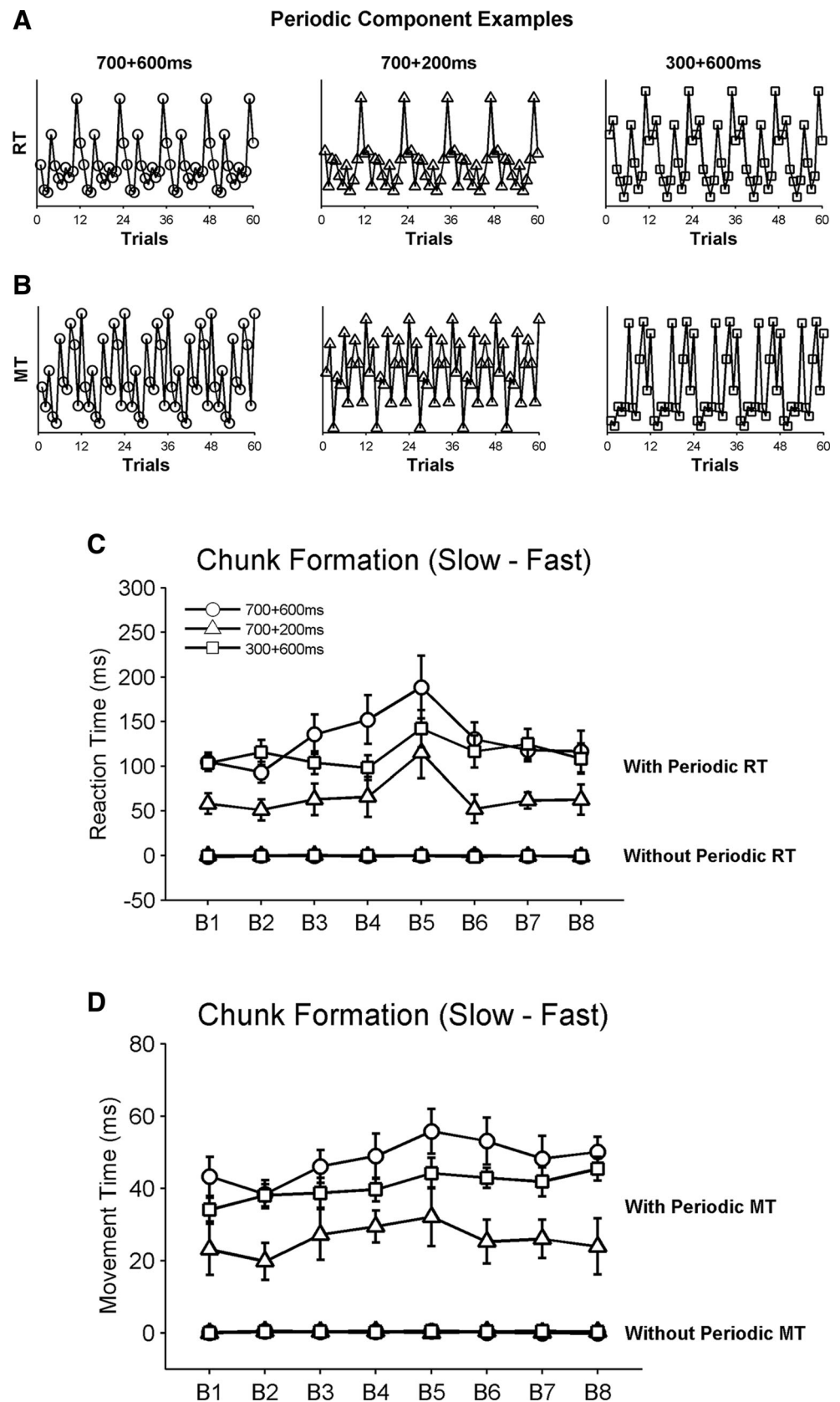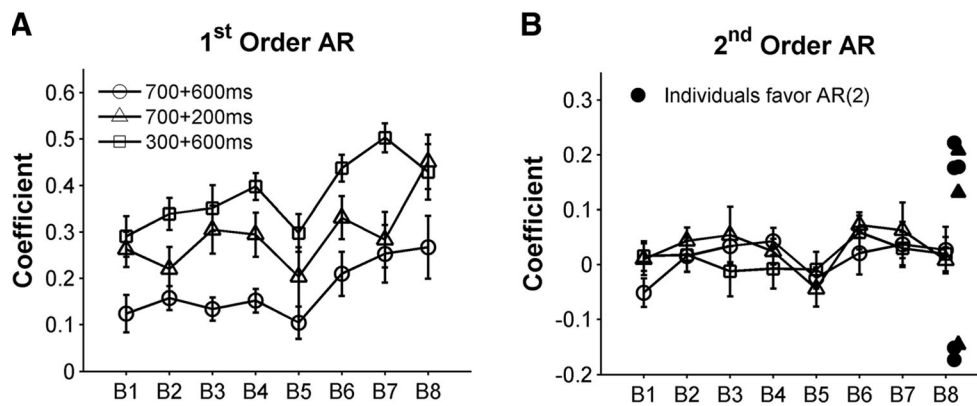
Fig. 4 **a** Mean coefficient of the first-order autoregressive term. **b** Mean coefficient of the second-order autoregressive term and the coefficient of eight individuals whose RT performance prefers AR (2) model. *Error bars* represent standard errors

transition). These results taken together suggest that the early acquisition of implicit sequences is very likely to result from statistical learning rather than chunk learning.

Chunking is suggested as a core mechanism underlying sequence learning (Gobet et al., 2001). Chunking benefits sequence learning as it attenuates memory loads during learning by segmenting a long sequence into shorter segments (Bo & Seidler, 2009; Wymbs, Bassett, Mucha, Porter, & Grafton, 2012). These shorter segments are normally indicated by a slower response followed by a few quicker responses (Bo & Seidler, 2009; Koch & Hoffmann, 2000). In these studies, slower responses were identified at certain positions in a sequence. However, our data imply that chunks identified in such a way may result from physical constraints of responses, such as transitions between feet and response locations. For example, an individual reacted and moved slower to targets at some locations (e.g., location 1 vs. 2) or locations corresponding to certain effectors (e.g., left vs. right foot). The same slow versus fast RT and MT patterns repeated when the sequence recurred, forming a typical chunk pattern where a slower response was followed by faster response(s). When the component of RT and MT originating from the biomechanical effects was taken into consideration, chunking was no longer observed. Although this study employed a foot-stepping SRT task in which biomechanical effects would be greater compared with a classic finger-pressing SRT task, we surmise that the same effects of biomechanical constraints on chunk learning is very likely to be true in a finger-pressing SRT task. For example, we suspect that the hand or finger transition pattern and/or various fingers would lead to different response speeds. These biomechanical effects on finger response speed could yield illusory chunk formations. To examine our hypothesis, future studies using the classic SRT task are needed.

Our results challenge the chunking hypothesis for learning implicit sequences—results that are consistent with recently reported findings. For example, it has been demonstrated by our data as well as previous studies that chunks began to form as soon as an individual starts the SRT task, but learning took place

without improvements in these chunks (Jimenez, 2008; Song & Cohen, 2014). In addition, the chunking principle is contradicted with memory consolidation process during sequence learning. Specifically, there is no time allowed for memory of each chunk to stabilize since chunks are performed continuously. Consequently, chunks would interfere with each other and thus prevent learning of the whole sequence (Robertson, Pascual-Leone, & Miall, 2004). This converging evidence indicates that chunking appears not to be the mechanism underlying the initial acquisition of implicit motor sequences.

An alternative to chunk learning arises from the learning of statistical structure underlying the sequence (Bornstein & Daw, 2012; Hunt & Aslin, 2001). Statistical learning prevails in the cognitive learning literature (see, Perruchet & Pacton, 2006, for a review). Our results favor the statistical learning interpretation. We found that each RT depended on the preceding RT, as revealed by the first-order autocorrelation.[1] The serial dependence is ubiquitous in human cognitive performance and has been suggested to reflect an internal representational structure within the cognitive system (Farrell, Wagenmakers, & Ratcliff, 2006; Gilden, Thornton, & Mallon, 1995). In particular, this self-dependency in RT implies that there is an association between the memory state of the previous and current stimulus (Verstynen et al., 2012). Importantly, given that the magnitudes of first-order autoregressions increased as practice progressed, the first-order autocorrelations are more likely to result from learning rather than preexisting cognitive dynamics. Meanwhile, this trial-to-trial dependency was not under the influence of biomechanical constraints on foot stepping, because 1) the biomechanical effects were explained by a periodic component in our model,

---

[1] In addition to the first-order autoregression, second-order autoregressions were found in the RTs of a few participants at the end of learning, especially those who performed the SRT task under a long ISI condition. Although future studies are necessary, this result suggests that as learning progresses, the higher-order statistical structure (i.e., statistical transitions between non-adjacent elements) may also be acquired.

and 2) MT would exhibit a similar serial dependency under the effect of biomechanical constraints but our results found no self-dependency in MT. One caveat worthy of consideration is that we did not find statistical differences in the first-order autoregression between blocks 4 and 5, although there was a trend that sequence B in block 5 interrupted the serial dependency (Fig. 4a). This is presumably because sequences A and B shared fewer similar transitional properties. For example, 1 was followed by 4 with 50 % chance in both sequences. In future studies, we would suggest that two completely distinguishable sequences be used.

Although our results do not favor chunk learning, there is no doubt that the chunking principle is ubiquitous in human cognitive and motor skills. The key conclusion of statistical but not chunk learning made here is specifically regarding the *initial* acquisition of *implicit* motor sequences in a SRT task under *fixed ISI* conditions. It is possible that the use of chunk or statistical learning is sensitive to various factors in the learning of sequences. The first factor to consider is the timescale of learning. Chunk formations usually develop after statistical learning, because they are built based on the stimulus contingencies acquired through statistical learning (Fiser & Aslin, 2005). With sequence chunks, motor skills could be executed automatically, benefitting the speed, accuracy, and cognitive demand of motor planning (Rosenbaum, Kenny, & Derr, 1983; Sakai, Hikosaka, & Nakamura, 2004). Thus, chunking may not develop in the early stages of learning. Rather, it serves as a hierarchical representation in later stages after the sequence structure is learned. The second potential factor involves explicit and implicit memory. There was a parallel between statistical learning and implicit memory (Meulemans & Van der Linden, 2003; Perruchet & Pacton, 2006) as well as between chunk learning and explicit memory (Curran, 1995; Koch, Philipp, & Gade, 2006; Meulemans & Van der Linden, 2003; but see Boyd, et al., 2009, for different results). Because sequence learning observed in this study is more implicit than explicit (Supplementary Methods and Results), statistical learning is rather likely to be used.

The third factor involves the task pacing. Chunking has been shown often to be accompanied by rhythmic movement (i.e., slow-fast-fast) (Sakai, et al., 2004). This raises the possibility of chunk learning in a self-paced SRT task (Cohen, et al., 1990; Curran & Keele, 1993; Koch & Hoffmann, 2000). In the self-paced SRT task, the response-stimulus interval (RSI) rather than the inter-stimulus interval (ISI) is controlled. The learner could flexibly prolong the ISI preceding a chunk and shorten the ISI(s) within the chunk, yielding an adjustable ISI and thus rhythmic movement. However, the fixed ISI used in our study may prevent the emergence of such rhythmic performance and thus chunk learning. To further elucidate the role of statistical and chunk learning, future studies are necessary to address these variations in the SRT task.

Conclusions

We found that sequence learning takes place under different stimulus interval conditions. Importantly, we found that the initial acquisition of implicit sequences may arise from statistical learning rather than chunk learning. Although chunk formations also were observed, we demonstrated that these chunks resulted from preexisting biomechanical constraints. We suggest that task constraints that could impact the emergence of statistical and chunk learning should be further investigated in future studies.

References

Bo, J., & Seidler, R. D. (2009). Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *Journal of Neurophysiology, 101,* 3116–3125.

Bornstein, A. M., & Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience, 35,* 1011–1023.

Boyd, L. A., Edwards, J. D., Siengsukon, C. S., Vidoni, E. D., Wessel, B. D., & Linsdell, M. A. (2009). Motor sequence chunking is impaired by basal ganglia stroke. *Neurobiology of Learning and Memory, 92,* 35–44.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology-Learning Memory and Cognition, 16,* 17–30.

Curran, T. (1995). On the neural mechanisms of sequence learning. *Psyche, 2.*

Curran, T., & Keele, S. W. (1993). Attentional and nonattentional forms of sequence learning. *Journal of Experimental Psychology-Learning Memory and Cognition, 19,* 189–202.

Destrebecqz, A., & Cleeremans, A. (2003). Temporal effects in sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 181–213). Amsterdam: John Benjamins Publishing Company.

Farrell, S., Wagenmakers, E. J., & Ratcliff, R. (2006). 1/f noise in human cognition: Is it ubiquitous, and what does it mean? *Psychonomic Bulletin & Review, 13,* 737–741.

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General, 134,* 521.

Frensch, P. A., & Miner, C. S. (1994). Effects of presentation rate and individual differences in short-term memory capacity on an indirect measure of serial learning. *Memory & Cognition, 22,* 95–110.

Gilden, D. L., Thornton, T., & Mallon, M. W. (1995). 1/f noise in human cognition. *Science, 267,* 1837–1839.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences, 5,* 236–243.

Gobet, F., Lloyd-Kelly, M., & Lane, P. C. R. (2016). What's in a name? The multiple meanings of "chunk" and "chunking". *Frontiers in Psychology 7.*

Hunt, R., & Aslin, R. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology-General, 130,* 658–680.

Jimenez, L. (2008). Taking patterns for chunks: Is there any evidence of chunk learning in continuous serial reaction-time tasks? *Psychological Research, 72,* 387–396.

Kirsch, W., Sebald, A., & Hoffmann, J. (2010). RT patterns and chunks in SRT tasks: A reply to Jimenez (2008). *Psychological Research, 74,* 352–358.

Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychological Research, 63,* 22–35.

Koch, I., Philipp, A. M., & Gade, M. (2006). Chunking in task sequences modulates task inhibition. *Psychological Science, 17,* 346–350.

Meulemans, T., & Van der Linden, M. (2003). Implicit learning of complex information in amnesia. *Brain and Cognition, 52,* 250–257.

Moisello, C., Crupi, D., Tunik, E., Quartarone, A., Bove, M., Tononi, G., & Ghilardi, M. F. (2009). The serial reaction time task revisited: a study on motor sequence learning with an arm-reaching task. *Experimental Brain Research, 194,* 143–155.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19,* 1–32.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238.

Robertson, E. M. (2007). The serial reaction time task: Implicit motor skill learning? *Journal of Neuroscience, 27,* 10073–10075.

Robertson, E. M., Pascual-Leone, A., & Miall, R. C. (2004). Current concepts in procedural consolidation. *Nature Reviews Neuroscience, 5,* 576–582.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology-Human Perception and Performance, 9,* 86–102.

Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences, 8,* 547–553.

Song, S., & Cohen, L. (2014). Impact of conscious intent on chunking during motor learning. *Learning & Memory, 21,* 449–451.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57,* 421–457.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97,* B25–B34.

Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology-General, 134,* 552–564.

Verstynen, T., Phillips, J., Braun, E., Workman, B., Schunn, C., & Schneider, W. (2012). Dynamic sensorimotor planning during long-term sequence learning: The role of variability, response chunking and planning errors. *PLoS One, 7*(10), e47336.

Visser, I., Raijmakers, M. E. J., & Molenaar, P. C. M. (2007). Characterizing sequence knowledge using online measures and hidden Markov models. *Memory and Cognition, 35,* 1502–1517.

Willingham, D. B., Greenberg, A. R., & Thomas, R. C. (1997). Response-to-stimulus interval does not affect implicit motor sequence learning, but does affect performance. *Memory and Cognition, 25,* 534–542.

Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron, 74,* 936–946.