

# The Evolution of Chunks in Sequence Learning

Laure Tosatto, Joël Fagot, Dezso Nemeth, Arnaud Rey

▶ **To cite this version:**

Laure Tosatto, Joël Fagot, Dezso Nemeth, Arnaud Rey. The Evolution of Chunks in Sequence Learning. Cognitive Science, 2022, 46 (4), pp.e13124. 10.1111/cogs.13124 . hal-03640218

**HAL Id: hal-03640218**
**https://hal.science/hal-03640218**

Submitted on 13 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The Evolution of Chunks in Sequence Learning

Laure Tosatto[a,b], Joël Fagot[a,b,c], Dezso Nemeth[d,e,f], and Arnaud Rey[a,b]

a. Aix Marseille Univ, CNRS, LPC, Marseille, France
b. Aix Marseille Univ, ILCB, Aix-en-Provence, France
c. Station de Primatologie - Celphedia, CNRS UAR846, Rousset, France
d. Université Claude Bernard Lyon 1, Lyon Neuroscience Research Center, Lyon, France
e. Institute of Psychology, Eotvos Lorand University, Budapest, Hungary
f. Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Budapest, Hungary

Running head: EVOLUTION OF CHUNKS

Correspondence concerning this article should be addressed to Laure Tosatto, Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Université, 3, place Victor Hugo, Case D, 13331 Marseille Cedex 03, France. Email: laure.tosatto@univ-amu.fr.

**Abstract**

Chunking mechanisms are central to several cognitive processes and notably to the acquisition of visuo-motor sequences. Individuals segment sequences into chunks of items to perform visuo-motor tasks more fluidly, rapidly, and accurately. However, the exact dynamics of chunking processes in the case of extended practice remain unclear. Using an operant conditioning device, eighteen Guinea baboons (*Papio papio*) produced a fixed sequence of nine movements during 1,000 trials by pointing to a moving target on a touch screen. Response times analyses revealed a specific chunking pattern of the sequence for each baboon. More importantly, we found that these patterns evolved during the course of the experiment, with chunks becoming progressively fewer and longer. We identified two chunk reorganization mechanisms: the *recombination* of preexisting chunks and the *concatenation* of two distinct chunks into a single one. These results provide new evidence on chunking mechanisms in sequence learning and challenge current models of associative and statistical learning.

*Keywords:* chunking, sequence learning, statistical learning, associative learning, non-human primate

## 1. Introduction

A key mechanism allowing our cognitive system to compress information and increase short term memory capacity is the formation of chunks (Mathy & Feldman, 2012; Miller, 1956). Chunking is defined as the process of associating and grouping several items together into a single processing unit (Gobet & al, 2001; Gobet, Lloyd-Kelly and Lane, 2016). In coherence with the limits encountered by our cognitive system (Cowan, 1988, 2017), the storage capacity of the chunks themselves seems limited (3 or 4 items per chunk, Allen & Coyne, 1988; Chase & Simon, 1973; Johnson, 1970). In the field of perceptual-motor learning, chunking has been considered as the main motor sequence integration mechanism (Diedrichsen & Kornysheva, 2015; Wymbs et al., 2012).

Perceptual-motor sequence learning is commonly described as the process by which a sequence of movements is acquired and executed with increased speed and accuracy (Willingham, 1998). Current sequence learning paradigms use a sequential button-press task: participants are presented with a series of visual stimuli organized in a sequence and are asked to press a corresponding response button (e.g., Cohen, Ivry, & Keele, 1990; Nemeth et al., 2010; Verwey, 2001). A faster and more accurate performance over time reflects learning of the sequence. Typically, researchers have reported robust effects of sequence's length on participants' response times (RTs): for short sequences (<4 keypresses), a decrease in the successive RTs can be observed when the sequence is learned; for longer sequences though, an initial decrease in RTs for the first 3 to 4 keypresses is followed by a longer RT on the next position. Then, RTs start decreasing again for the next 3 to 4 keypresses (e.g., Bo & Seidler, 2009; Verwey & Eikelboom, 2003). Similar results have been observed in non-human individuals (Terrace, 2002; Scarf et al., 2018). This phenomenon has been interpreted as the long sequence being spontaneously segmented into shorter motor chunks reflecting the

sequence organization in memory (Sakai et al., 2003). Long temporal gaps between responses are assumed to mark chunk boundaries (Abrahamse et al., 2013; Bottary et al., 2016).

In the case of extended practice, several studies have investigated the evolution of chunks and have reported a decrease over time of the number of chunks (e.g., Acuna et al., 2015; Bera et al., 2021; Ramkumar et al., 2016). One mechanism proposed to account for this phenomenon is the *concatenation* of chunks (Verwey, 2001; Wright et al., 2010). Concatenation is described as independent chunks being executed more fluidly with practice and with a decrease of the temporal gap between them leading to a single and longer chunk (Abrahamse et al., 2013). However, the precise nature of the mechanisms allowing RTs to decrease and independent chunks to merge into a larger functional unit remain unclear. The present study aims at collecting new evidence on the mechanisms underlying the evolution of chunking mechanisms in a perceptual-motor sequence learning task.

One characteristic of previous sequence learning studies is the pattern of transition probabilities (TPs) between items of the sequence. In an experimental set up where only a few response buttons are available to the participant (e.g., Grafton, et al., 2002; Verwey, 2001; Willingham, 1999), some stimuli are necessarily presented multiple times within the same sequence. For instance, in a sequence such as A-B-A-C-D-B-C, A is either followed by B or C, therefore the probability of B given A is 0.5, whereas D is always followed by B thus the probability of B given D is 1. This heterogeneity of TPs may constrain the strength of the connections formed between the successive elements of the sequence and affect the resulting chunking pattern (Aslin, 2017; Lazartigues et al. 2021).

Another characteristic of previous studies is, apart from a few exceptions (e.g., Conway & Christiansen, 2001), that most sequence learning studies have been conducted with human participants. However, even if implicit instructions are provided, humans may use their inner language and develop strategies to perform the task (Rey et al., 2019). This may affect both

their performance and their chunking processes, and it is hard to tease apart these verbal and explicit influences from associative and chunking mechanisms.

To address these two concerns, we used the serial pointing task developed by Minier et al. (2016). In this task, participating monkeys had to touch a target red circle that could appear at 9 possible locations, defined by a 3 x 3 matrix of crosses displayed on a touch screen. They were repeatedly exposed to the same sequence of 9 locations that always appeared in the same order upon a thousand trials. The repeated sequence was composed of the nine possible locations presented in a specific order (e.g., 9-5-6-8-2-4-7-1-3) and leading to a transitional probability between locations equal to 1. RTs for each location were recorded, providing us with detailed information about the temporal dynamics of sequence processing. This extended practice on a long single repeated sequence allowed us to collect new evidence on the formation and the evolution of chunks over time.

## 2. Method

### *2.1. Participants*

Thirteen female and five male Guinea baboons (*Papio papio*, age range 2.8—23.7 years) from the CNRS primate facility in Rousset (France), living in a social group of 25 individuals, were tested in this study. Water was provided *ad libitum* during the test, and the monkeys received their normal food ration of fruits every day at 5 PM. For practical reasons, we stopped the experiment after 18 monkeys completed all scheduled trials.

### *2.2. Materials*

#### *2.2.1. Apparatus*

This experiment was conducted with a computer-learning device based on the voluntary participation of baboons (for details, see Fagot & Bonté, 2010). Baboons implanted with a RFID microchip had free access to 10 automatic operant conditioning learning devices equipped with touch screens. Each time a monkey entered a test chamber, it was identified by its microchip,

and the system resumed the trial list where the monkey left it at its previous visit. The experiment was controlled by E-prime (Version 2.0, Psychology Software Tools, Pittsburgh, PA, USA).

*2.2.2. Task and stimuli*

The screen was divided into nine equidistant predetermined locations represented by white crosses on a black background, virtually labeled as position 1 to 9 (see Figure 1). A trial began with the presentation of a yellow fixation cross at the bottom of the screen. Once pressed, the fixation cross disappeared and the nine white crosses were displayed, one of them being replaced by the target, a red circle. When the target was touched, it was immediately replaced by the cross. The red circle then replaced the next position in the sequence until it was touched, and a new position was displayed. Reward was provided at the end of a sequence of nine touches.

< insert Figure 1 here >

To learn the task, baboons initially received random trials that were rewarded after three touches. Then, the number of touches in a trial was progressively increased up to nine. If baboons touched an inappropriate location (incorrect trial) or failed to touch the screen within 5,000ms after the red circle's appearance (aborted trial), a green screen was displayed for 3,000ms as a marker of failure. Aborted trials were not retained and therefore presented again, while incorrect trials were not. The time elapsed between the appearance of the red circle and the baboon's touch on this circle was recorded as the RT.

*2.2.3. Design of the sequences*

To control the motor difficulty of the transitions to be produced in the sequence, a random phase of sequence production was first conducted, where thirteen baboons performed random sequences of six positions for 1,000 trials. Based on these random trials, a baseline measure for all possible transitions from one location to another was computed by calculating

mean RTs for each transition (e.g., from position 1 to 9), leading to a 9x9 matrix of mean RTs calculated over the entire group of baboons (see Appendix A). This baseline measure corresponds to the mean time baboons need to move their hand from one location to another, when the transition is unpredictable (i.e., random).

Based on these baseline measures, we designed two sequences of nine serial locations for which each transition T was faster (or equally fast) to produce than the next one (i.e., T1≤T2≤…≤T8; see Appendix B). This way, a decrease in RT for a given transition can be interpreted as the anticipation (or learning) of that position from the previous one.

### *2.3. Procedure*

Baboons were either presented with Sequence 1 (N=8 baboons) or Sequence 2 (N=10 baboons) and had to produce it for 1000 successive trials. RTs for each position of the sequence were recorded for all the trials.

## 3. Results

On average, baboons required 2.9 days to complete the 1,000 trials, with a mean of 339.7 trials per day and a mean accuracy level of 92.5% (*SD*=6.2%). Incorrect trials were removed from the dataset (7.8%). A recursive trimming procedure excluded RTs greater or smaller than 2.5 standard deviations from the participant's mean for each of the nine possible positions (24.4%)[1]. RTs for each of the nine positions and for the 1,000 trials were divided into 10 Blocks of 100 trials.

### *3.1. General sequence learning*

[1] Inspection of the response times distribution revealed that a majority of responses were produced around 500ms. A smaller group of RTs appeared around 1,000 ms and was likely due to situations in which baboon's response was not recorded by the computer, because their hands were dirty. In this situation, they had to touch the screen again, and longer RTs were recorded (that are on average twice longer compared to the first responses). This is why we have adopted this recursive trimming procedure. After applying this procedure, there was still a mean number of 77.9 remaining RTs for each position per participant and per block of 100 trials (the minimum number of remaining RTs for one position being 32).

Learning of the sequence was estimated on mean RTs by a repeated measures one-way ANOVA with Block (1 to 10) as the within factor. The effect of Block was highly significant (Block 1, $M$=452.8, $SD$=45.3; Block 10, $M$=400.1, $SD$=56.3), $F(1,9)=30.43$, $p<.001$, $\eta^2=0.642$, indicating that mean RTs decreased throughout the blocks of trials and that monkeys learned the sequence.

*3.2. Chunking of the sequence*

To study the chunking pattern of the sequence by monkeys, we adopted the following reasoning. We considered successive positions *A* and *B* to be part of the same chunk as long as the transition time from one position to the next did not correspond to a significant increase in RT. Therefore, if $RT_A<RT_B$ then *A* and *B* do not belong to the same chunk, and the AB transition marks a chunk boundary (Kennerly et al., 2004). This reasoning was applied on each Block and for each monkey. Statistical significance was assessed through paired-sample t-tests for each pair of successive positions. Each time the RT of a pair's second position was significantly higher than the first position, it marked a chunk boundary. Figure 2 illustrates this method on the performances of one monkey.

< insert Figure 2 here >

*3.3. Evolution of chunks*

With this method, we were able to quantify the number of chunks and their average size produced on each block by each monkey. For example, the monkey from Figure 2 produced 3 chunks and an average chunk size of 3 in Block 1. In Block 10, the number of chunks was 2 and the average chunk size was 4.5. Figure 3 reports the average number of chunks and chunk size for all monkeys across the 10 blocks of the experiment. For example, we found that the mean chunk size was 2.2 positions (*CI*=0.31; *Min*=1; *Max*=5) in the first block and 3.375 (*CI*=0.53; *Min*=1; *Max*=8) in the last block. Two repeated measures one-way ANOVA were conducted to test the effect of Block on the mean number of chunks and the average chunk size.

This analysis revealed that the number of chunks significantly decreased across blocks, $F(1,9)=9.421$, $p<.001$, $\eta^2=.357$, and that the average chunk size significantly increased across blocks, $F(1,9)=4.794$, $p<.01$, $\eta^2=.22$.

The evolution of these two indicators can be accounted for by two reorganization mechanisms: the *recombination* of pre-existing chunks (50.4% of the reorganizations, observed for 15 monkeys) and the *concatenation* of two chunks into a single one (49.6% of the reorganizations, observed for 17 monkeys). An illustration of these reorganizations is provided in Figure 2: from Block 1 to Block 3, the first six positions of the sequence were parsed into two equivalent chunks of three positions, indicated by the significant increase in RTs between Position 3 and 4. Starting Block 4, the chunks are recombined with Position 4 being slowly integrated to the first chunk as its difference with Position 3 disappears and a significant increase in RTs is now appearing between Position 4 and 5. The concatenation of two chunks into a single one is also illustrated in Figure 2: while Chunks 2 and 3 were separated by a significant increase in RT between Position 6 and Position 7 until Block 8, this difference disappeared in Block 9, and the two chunks were then grouped into one. Table 1 provides a detailed description of the chunks appearing in Figure 2. For each block of trials, it counts the number of chunks, their respective size, and each occurrence of a concatenation or a recombination. At the group level, Table 2 provides the total number of concatenations and recombinations obtained for each block and for all monkeys[2]. A repeated measure ANOVA with Block and Mechanisms (concatenation vs. recombination) did not reveal any significant effect (all ps > .05).

< insert Table 1 & 2 here >

### *3.4. Additional analyses*

[2] See Supplementary material for a description of the patterns of responses for each monkey over blocks. The same information as the one displayed in Figure 2 and Table 1 for the monkey "Atmosphere" is provided for all monkeys from the present experiment (see Figures and Tables S1 to S17). These data clearly show that each monkey displays a specific chunking pattern that follows a specific evolution over blocks of trials.

Our data show that a chunking pattern emerges and evolves over time for each monkey (i.e., Figures S1 to S17). The averaged data on Block 1 indicates that a first chunking pattern rapidly emerges after a few dozen of trials. However, we checked that monkeys produced the same ascending pattern of RTs during their first 10 trials on the repeated sequence than the one we computed from the random baseline trials (see Figure S18). This ensures that our sequences were correctly designed and that chunking mechanisms are rapidly efficient.

We also compared the mean response times computed on every trial and averaged over blocks in the random and repeated sequence conditions. The result indicates that chunking mechanisms qualitatively and quantitatively influenced the monkey's behavior by significantly increasing their global efficacy in this visuo-motor task (see Figure S19).

**4. Discussion**

Three main findings were obtained in the present study. First, we found that non-human primates spontaneously segmented long sequences into short chunks. Second, with extended practice, chunks became longer and fewer. Third, based on these observations, we found this decrease in the number of chunks and this increase in chunks' size was due to two types of reorganizations: the recombination of several preexisting chunks and the concatenation of two distinct chunks into a single one.

Our first finding is consistent with previous studies on sequence learning in humans as participants were found to initially segment the sequence into small chunks (e.g., Nissen & Bullemer, 1987; Verwey, 1996; 2001; 2003, Verwey et al., 2002). Chunks typically contained 2 to 4 items, sometimes 5 (Sakai et al., 2003), which is similar to what we observed in the first block of the experiment. With extended practice, the mean chunk size could reach up to 8 successive positions. However, these large chunks are very rare in our data and in the literature, although Kennerly, Sakai and Rushworth (2004) found in humans a mean chunk size of 7.83 items for sequences of 12 items. This difference is likely due to motor constrains that vary

238 across experimental paradigms and may facilitate the development of longer chunks. This first
239 result indicates that chunking is a fundamental mechanism of sequence learning in both human
240 and non-human primates, and that the cognitive system spontaneously forms strong
241 associations between repeated co-occurring events (e.g., Perruchet & Vinter, 1998; 2002; Rey
242 et al., 2019). However, the results also show that there is an initial limit to the number of
243 associations that can be formed successively.

244 Our second finding is in line with previous research and indicates that extended practice
245 allows the baboons to exceed these limits, as we found that chunks can be reorganized into
246 longer segments. This feature is robust and we observed reorganizations in 17 monkeys out of
247 18. Ramkumar et al. (2016) suggested that to limit the cost of computation, learning new
248 sequences of movements starts with many short chunks. With practice, the execution of short
249 chunks becomes more efficient, which reduces the computation's complexity. This increase in
250 efficiency for short chunks would promote more complex computations leading to the
251 development of longer chunks. However, the precise neural mechanisms supporting these
252 changes over time are still unclear, as are the conditions that promote switching from one
253 pattern to another.

254 Our third main finding is related to the evolution of chunks and the two reorganization
255 mechanisms that allow chunks to become longer: the concatenation and recombination of
256 chunks. Abrahamse et al. (2013) define the concatenation mechanism as the process by which
257 two successive chunks are performed more fluidly and the temporal gap between them
258 decreases. This description is consistent with some of our findings, but concatenations were not
259 the only form of reorganization we observed. To our knowledge, there is no other study
260 reporting the second type of reorganization we found, i.e., recombinations. Indeed,
261 recombinations differs from concatenations by corresponding to the emergence of a new
262 segmentation pattern across chunks. As shown in Table 2, the occurrence of a concatenation or

a recombination does not seem related to a specific stage of learning as it seems to be evenly distributed throughout the experiment.

The absence of differences in the distribution of recombinations and concatenations throughout learning can be surprising considering that, in earlier stages of learning, chunks are less stable, and one would expect that recombinations would occur more frequently at the beginning of the task. Conversely, as the sequence is learned, chunks stability should increase, and one would expect more concatenations to occur. However, as indicated by the linear decrease in mean RTs (see Figure S19), it is possible that learning of the sequence was not over even after 1,000 trials, and that baboons did not yet reached their optimal performance. Therefore, an increase of the number of concatenations over the number of recombinations could occur during later stages of learning that were not observed here.

The present findings therefore indicate that baboons initially organize their pattern of responses by segmenting the sequence of nine touches into small chunks (i.e., from 2 to 4 elements). With extended practice, the chunking pattern either remain the same or is reorganized be either concatenating or recombining chunks. We identify three parameters that are likely determining the dynamical evolution of a chunking pattern. The first parameter corresponds to the *internal stability* of chunks. Consider, for example, Chunk 3 in Figure 2 from Block 2 to 4. The mean RTs for Position 7, 8 and 9 of Chunk 3 are extremely far apart, indicating that this chunking pattern is highly stable. To operationalize this notion of stability, one could compute the standard deviation among these three RTs and find a large value (e.g., the standard deviation in Block 2 for Chunk 3 is 68.8). Alternatively, Chunk 1 in Block 3 certainly has a weaker internal stability due to the small differences between the five RTs composing this chunk (the standard deviation is here equal to 19.7). Internal stability of chunks therefore likely determines the survival of a chunking pattern or its reorganization.

The second parameter corresponds to the *stability of chunk boundaries* that can be estimated by computing the difference between the last RT of one chunk and the first RT of the next chunk. Consider, for example, the boundary between Chunk 2 and 3 in Block 2 of Figure 2, the difference between the last RT of Chunk 2 (i.e., 378.5 ms) and the first RT of Chunk 3 (i.e., 501.1 ms) is equal to 122.6 ms. Here the difference is huge, and this boundary does not change until Block 8.

The third parameter is related to the *increased efficiency* in the realization of chunks during extended practice (an idea proposed by Ramkumar et al., 2016). Indeed, the reason why the boundary between Chunk 2 and 3 decreases from Block 1 to 7 is likely due the increase in efficiency in the realization of Chunk 3 (leading to faster RTs) and consequently, to the smaller difference between the last RT of Chunk 2 and the first RT of Chunk 3.

The present data therefore suggest that the evolution of chunks in sequence learning probably depends on these three parameters: 1) the internal stability of chunks; 2) the stability of chunk boundaries; and 3) the increased efficiency in the realization of chunks during extended practice. Concatenations will occur when two chunks have a strong internal stability and when the realization of the second chunk will become more efficient, leading to a lower difference between the last RT of the first chunk and the first RT of the second chunk and therefore, a lower stability of chunks boundary. Recombinations are likely to be observed when the internal stability of chunks is low and when there is a low stability of chunks boundary. Finally, the chunking pattern is likely to remain the same when the chunk stability and the chunk boundaries are high.

Note that the fact that some chunks are being recombined throughout the task also challenges the idea that there is always a hierarchical organization in sequence learning (Restle, 1970; Yokoi & Diedrichsen, 2019) and that the reorganizations only operate by concatenating

311 preexisting chunks. Indeed, chunks appear to be more flexible than previously thought and
312 linear associations between adjacent elements does not seem to be the only mechanism at work.
313 Interestingly, current models of chunking only partially account for the reorganizations
314 we found. Most models, like the Competitive Chunking model (Servan-Schreiber & Anderson,
315 1990) or the PARSER model (Perruchet & Vinter, 1998) assume that repeated sequences lead
316 to the formation of chunks that are stored in memory. They also assume that two co-occurring
317 chunks can lead to a new and larger chunk by concatenating two previously smaller chunks.
318 Recombination, though, is not a mechanism implemented in any model of chunking. The
319 present data, therefore, represent a strong challenge for current models of statistical and
320 associative learning.
321 To summarize, these data indicate that when non-verbal non-human animals are
322 repeatedly exposed to a long sequence of 9 elements, associations are formed initially between
323 a limited number of co-occurring elements (i.e., from 2 to 4 elements). With extended practice,
324 these patterns of associations can become longer. The reorganization mechanisms supporting
325 these evolutions are not yet accounted for by any model of associative learning but
326 understanding the dynamics of these mechanisms represents a strong challenge for the future
327 generation of computational models of associative and statistical learning.
328
329

**Acknowledgements**

This work was supported by the BLRI Labex (ANR-11-LABX-0036), Institut Convergence ILCB (ANR-16-CONV-0002), the CHUNKED ANR project (#ANR-17-CE28-0013-02) and IDEXLYON Fellowship of the University of Lyon as part of the Programme Investissements d'Avenir (ANR-16-IDEX-0005). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We are grateful to Laura Ordonez-Magro for her helpful comments on a previous version of the manuscript.

**Open Practices Statements**

Data from the experiment are available on Open Science Framework at https://osf.io/xcw95/?view_only=73f8cc59abe343948a6c6ea090d6dbe2.

**Conflict of Interest**

The authors declare that they have no conflict of interest.

**Animal rights**

This research adhered to the applicable French rules for ethical treatment of research animals and received ethical approval from the French Ministry of Education (approval APAFIS#2717-2015111708173794 10 v3).

## References

Abrahamse, E., Ruitenberg, M., De Kleine, E., & Verwey, W. B. (2013). Control of automated behavior: insights from the discrete sequence production task. *Frontiers in Human Neuroscience*, *7*, 82.

Acuna, D. E., Wymbs, N. F., Reynolds, C. A., Picard, N., Turner, R. S., Strick, P. L., ... & Kording, K. P. (2014). Multifaceted aspects of chunking enable robust algorithms. *Journal of neurophysiology*, *112*(8), 1849-1856.

Allen, P. A., & Coyne, A. C. (1988). Age differences in primary organization or processing variability? Part I: An examination of age and primary organization. *Experimental Aging Research*, *14*(3), 143-149.

Aslin, R. N. (2017). Statistical learning: a powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2), e1373.

Bera, K., Shukla, A., & Bapi, R. S. (2021). Motor chunking in internally guided sequencing. *Brain Sciences*, *11*(3), 292.

Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory correlate with deficits in explicit motor sequence learning. *Journal of Neurophysiology*, *102*(5), 2744-2754.

Bottary, R., Sonni, A., Wright, D., & Spencer, R. M. (2016). Insufficient chunk concatenation may underlie changes in sleep-dependent consolidation of motor sequence learning in older adults. *Learning & Memory*, *23*(9), 455-459.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81.

Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(1), 17.

Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, *5*(12), 539-546.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*(2), 163.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158-1170.

Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and execution. *Trends in Cognitive Sciences*, *19*(4), 227-233.

Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio papio). *Behavior Research Methods*, *42*(2), 507-516.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236-243.

Gobet, F., Lloyd-Kelly, M., & Lane, P. C. (2016). What's in a name? The multiple meanings of "Chunk" and "Chunking". *Frontiers in Psychology*, *7*, 102.

Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the nondominant left hand. *Experimental Brain Research*, *146*(3), 369-378.

Johnson, N. F. (1970). The role of chunking and organization in the process of recall. In *Psychology of Learning and Motivation* (Vol. 4, pp. 171-247). Academic Press.

Kennerley, S. W., Sakai, K., & Rushworth, M. F. S. (2004). Organization of action sequences and the role of the pre-SMA. *Journal of Neurophysiology*, *91*(2), 978-993.

Lazartigues, L., Mathy, F., & Lavigne, F. (2021). Statistical learning of unbalanced exclusive-or temporal sequences in humans. *Plos one*, *16*(2), e0246826.

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*(3), 346-362.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Minier, L., Fagot, J., & Rey, A. (2016). The temporal dynamics of regularity extraction in non-human primates. *Cognitive Science*, *40*(4), 1019-1030.

Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental Brain Research*, *201*(2), 351-358.

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

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences*, *25*(3), 297-330.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, *77*(6), 481.

Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity extraction across species: Associative learning mechanisms shared by human and non-human primates. *Topics in Cognitive Science*, *11*(3), 573-586.

Ramkumar, P., Acuna, D. E., Berniker, M., Grafton, S. T., Turner, R. S., & Kording, K. P. (2016). Chunking as the result of an efficiency computation trade-off. *Nature Communications, 7(1)*, 1-11.

Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research*, *152*(2), 229-242.

Scarf, D., Smith, C. D., Jaswal, V. K., Magnuson, J. S., & Terrace, H. (2018). Chunky Monkey? The spontaneous temporal chunking of simultaneous chains by Humans (Homo Sapiens) and rhesus monkeys (Macaca mulatta). In A. Ibáñez-Contreras & B. Hernandez-Godinez (Eds), *Studies of Rhesus Monkeys and their Behaviors* (p. 39-58). Nova Science Publishers.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 592.

Song, S., & Cohen, L. (2014). Impact of conscious intent on chunking during motor learning. *Learning & Memory*, *21*(9), 449-451.

Terrace H.S. (2002) The Comparative Psychology of Chunking. In S.B. Fountain, M.D. Bunsey, J.H. Danks, & M.K. McBeath (eds), *Animal Cognition and Sequential Behavior* (p. 23-55). Springer, Boston, MA.

Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(3), 544.

Verwey, W. B. (2001). Concatenating familiar movement sequences: The versatile cognitive processor. *Acta Psychologica*, *106*(1-2), 69-95.

Verwey, W. B., & Eikelboom, T. (2003). Evidence for lasting sequence segmentation in the discrete sequence-production task. *Journal of Motor Behavior*, *35*(2), 171-181.

Verwey, W. B., Lammens, R., & van Honk, J. (2002). On the role of the SMA in the discrete sequence production task: a TMS study. *Neuropsychologia*, *40*(8), 1268-1276.

Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological Review*, *105*(3), 558.

Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory & Cognition*, *27*(3), 561-572.

Wright, D. L., Rhee, J. H., & Vaculin, A. (2010). Offline improvement during motor sequence learning is not restricted to developing motor chunks. *Journal of motor behavior*, *42*(5), 317-324.

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*(5), 936-946.

Yokoi, A., & Diedrichsen, J. (2019). Neural organization of hierarchical motor sequence representations in the human neocortex. *Neuron*, *103*(6), 1178-1190.

**Figure 1**

*Experimental display and stimuli presentation*

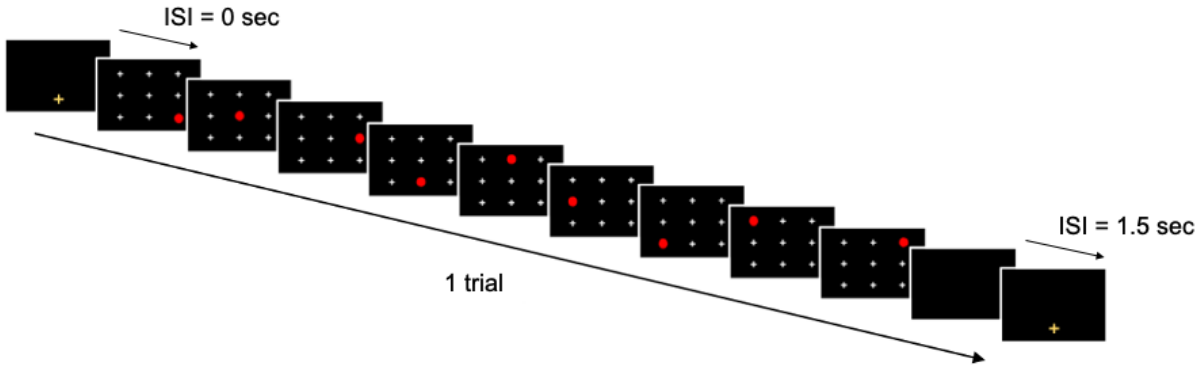
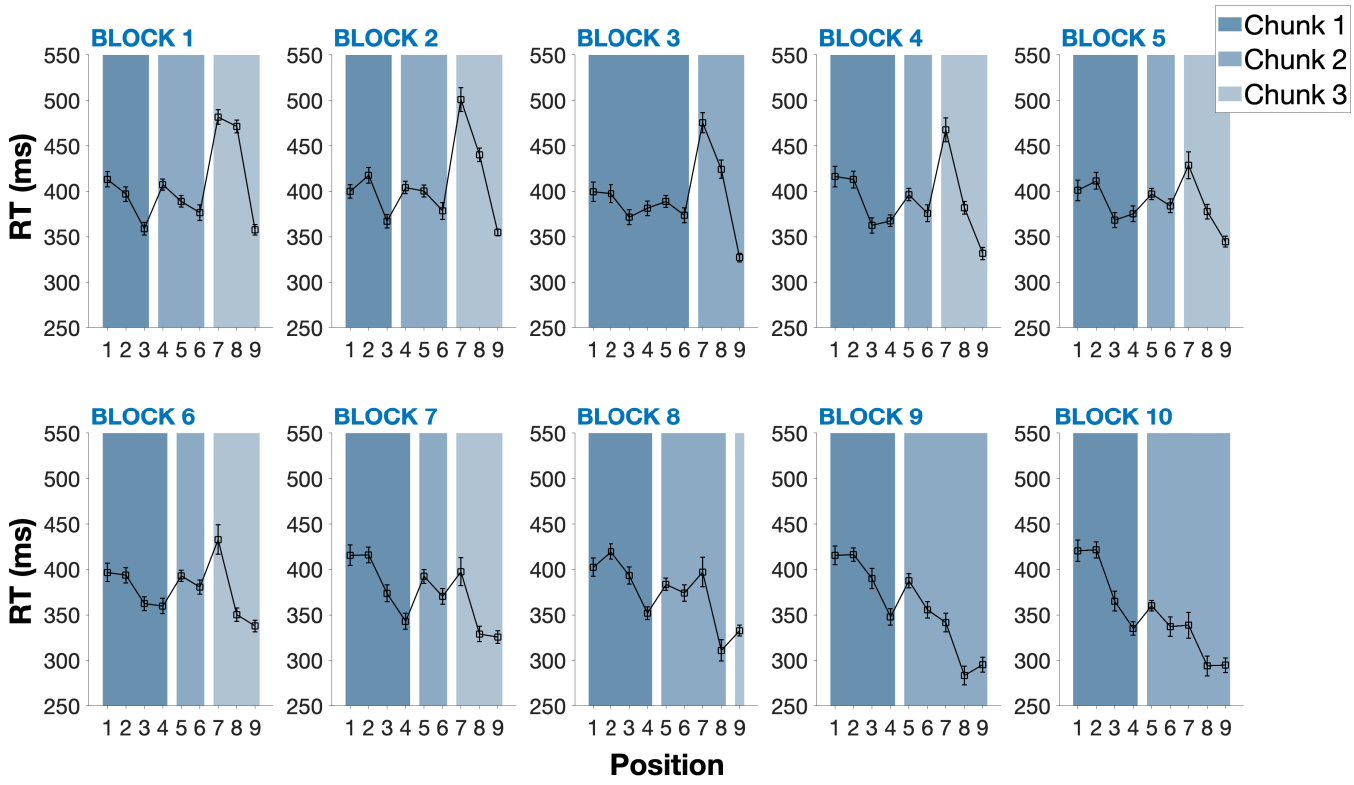


**Figure 2**

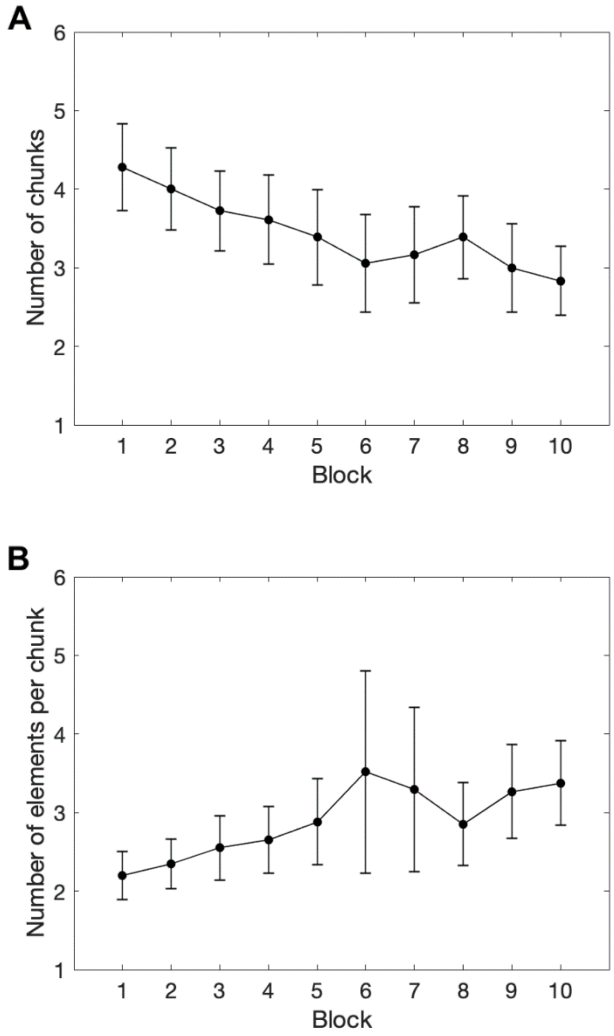
*Evolution of the chunking pattern for one individual (Atmosphere) throughout the task*




*Note.* Mean RT per position across the 10 blocks of trials for one baboon (Atmosphere) showing the evolution of the chunking pattern. This individual initially parses the sequence into three chunks of three positions in the first three blocks of trials. From Block 7 to 8, the 2nd and 3rd chunks (having a size of 2 and 3, respectively) are recombined into two chunks of size 4 and 1, respectively. From Block 8 to 9, the 2nd and 3rd chunks are concatenated into one single chunk of size 5. Error bars represent 95% confidence intervals.

**Figure 3**

*Evolution of chunks across blocks*




*Note*. A. Mean number of chunks per block. B. Mean chunk size (i.e. number of elements per chunk) per block (error bars represent 95%confidence intervals).

**Table 1**

*Summary table of the reorganizations observed throughout the task for one baboon (Atmophere).*

| | Chunk size | | | Nb of chunks | Concatenation | Recombinaison |
|---|---|---|---|---|---|---|
| | C1 | C2 | C3 | | | |
| Block 1 | 3 | 3 | 3 | 3 | - | - |
| 2 | 3 | 3 | 3 | 3 | - | - |
| 3 | 3 | 3 | 3 | 3 | - | - |
| 4 | 4 | 2 | 3 | 3 | - | 1 |
| 5 | 4 | 2 | 3 | 3 | - | 1 |
| 6 | 4 | 2 | 3 | 3 | - | - |
| 7 | 4 | 2 | 3 | 3 | - | - |
| 8 | 4 | 4 | 1 | 3 | - | 1 |
| 9 | 4 | 5 | - | 2 | 1 | - |
| 10 | 4 | 5 | - | 2 | - | - |
| Total | | | | | 1 | 3 |

*Note.* Numbers in columns 2-4 refer respectively to the size of the first (C1), second (C2) and third (C3) chunk (if any) in the sequence.

**Table 2**

*Total number of concatenations and recombinations per block.*

| Block | Concatenations | Recombinations |
|---|---|---|
| 2 | 11 | 9 |
| 3 | 8 | 7 |
| 4 | 6 | 7 |
| 5 | 6 | 4 |
| 6 | 8 | 5 |
| 7 | 3 | 6 |
| 8 | 1 | 10 |
| 9 | 8 | 3 |
| 10 | 5 | 6 |
| Total | 56 | 57 |

## Appendix A

*Mean response times over the entire group of baboons for each of the 72 possible transitions calculated from the 1000 random trials.*

| 1st position in Transition | 2nd position in Transition | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | - | 519 | 573 | 495 | 482 | 509 | 521 | 497 | 543 |
| 2 | 569 | - | 553 | 513 | 474 | 511 | 523 | 491 | 509 |
| 3 | 558 | 519 | - | 513 | 472 | 488 | 544 | 493 | 512 |
| 4 | 551 | 517 | 560 | - | 464 | 509 | 522 | 482 | 546 |
| 5 | 549 | 504 | 552 | 501 | - | 483 | 535 | 479 | 527 |
| 6 | 567 | 515 | 546 | 507 | 484 | - | 533 | 483 | 511 |
| 7 | 555 | 504 | 558 | 475 | 463 | 516 | - | 484 | 541 |
| 8 | 554 | 512 | 540 | 485 | 448 | 472 | 512 | - | 507 |
| 9 | 546 | 512 | 540 | 514 | 460 | 464 | 550 | 485 | - |

*Note.* All transitions are in milliseconds (ms) and correspond to the time elapsed between the disappearance of the red circle from the 1st position of the Transition and the monkey's touch on the 2nd position of the Transition.

## Appendix B

*Selected sequences and corresponding mean transition times (based on the baseline acquisition, see Appendix A)*

| Sequence | Position | | | | | | | | | Mean transition time (ms) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 |
| Seq. 1 (N=8) | 9 | 5 | 6 | 8 | 2 | 4 | 7 | 1 | 3 | 460 | 483 | 483 | 512 | 513 | 522 | 555 | 573 |
| Seq. 2 (N=10) | 7 | 5 | 6 | 8 | 2 | 4 | 9 | 1 | 3 | 463 | 483 | 483 | 512 | 513 | 522 | 555 | 573 |