Neuron

Article

Mental compression of spatial sequences in human working memory using numerical and geometrical primitives

Highlights

- The human brain compresses spatial sequences using an abstract language
- Language complexity modulates spatial anticipation signals
- Elementary geometrical operations can be decoded from MEG signals
- An ordinal number code reactivates periodically according to sequence structure

Authors

Fosca Al Roumi, Sébastien Marti, Liping Wang, Marie Amalric, Stanislas Dehaene

Correspondence

fosca.alroumi@gmail.com (F.A.R.), stanislas.dehaene@cea.fr (S.D.)

In brief

When humans see a sequence of spatial locations, their brains compress it in memory by using all available geometrical regularities. A hierarchical language of thought accounts for these observations: spatial, ordinal, and geometrical primitive codes can be extracted from brain activity.

Al Roumi et al., 2021, Neuron *109*, 1–13
August 18, 2021
https://doi.org/10.1016/j.neuron.2021.06.009

CellPress

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

Neuron

CellPress

Article

Mental compression of spatial sequences in human working memory using numerical and geometrical primitives

Fosca Al Roumi,[1,6,*] Sébastien Marti,[1,5] Liping Wang,[3] Marie Amalric,[1,4] and Stanislas Dehaene[1,2,*]
[1]Cognitive Neuroimaging Unit, CEA, INSERM, Université Paris-Saclay, NeuroSpin Center, 91191 Gif/Yvette, France
[2]Collège de France, Université Paris Sciences Lettres (PSL), 11 Place Marcelin Berthelot, 75005 Paris, France
[3]Institute of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China
[4]CAOS Lab, Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
[5]Deceased January 11, 2019
[6]Lead contact
*Correspondence: fosca.alroumi@gmail.com (F.A.R.), stanislas.dehaene@cea.fr (S.D.)
https://doi.org/10.1016/j.neuron.2021.06.009

SUMMARY

How does the human brain store sequences of spatial locations? We propose that each sequence is internally compressed using an abstract, language-like code that captures its numerical and geometrical regularities. We exposed participants to spatial sequences of fixed length but variable regularity while their brain activity was recorded using magneto-encephalography. Using multivariate decoders, each successive location could be decoded from brain signals, and upcoming locations were anticipated prior to their actual onset. Crucially, sequences with lower complexity, defined as the minimal description length provided by the formal language, led to lower error rates and to increased anticipations. Furthermore, neural codes specific to the numerical and geometrical primitives of the postulated language could be detected, both in isolation and within the sequences. These results suggest that the human brain detects sequence regularities at multiple nested levels and uses them to compress long sequences in working memory.

INTRODUCTION

Although non-human primates are able to learn sophisticated behavioral rules, the human species seems to be endowed with a deeper ability to discover the complex embedded structures that underlie the information present in the environment (Dehaene et al., 2015; Ferrigno et al., 2020; Fitch, 2004, 2014; Hauser et al., 2002; Jiang et al., 2018; Wang et al., 2015).

In the present research, which is part of series of behavioral and brain-imaging studies of spatial sequence learning (Amalric et al., 2017; Wang et al., 2019), we test the specific hypothesis that when they learn a spatial sequence, human subjects make use of a language-like system of nested rules of variable complexity. Imagine you need to remember a sequence of eight spatial locations. Remembering the sequence by storing each item in a memory slot would be difficult, as eight exceeds the typical working-memory span (Baddeley, 2003; Baddeley and Hitch, 1974; Botvinick and Watanabe, 2007; Hurlstone et al., 2014). However, if the first four items form a square and the next four draw another square, mentally compressing this sequence as ''two squares'' would facilitate its memorization. In the present work, we test whether working memory is organized as a flat set of slots or as a structured language.

In previous research (Amalric et al., 2017; Wang et al., 2019), we formalized the latter idea by proposing a hypothetical ''language of thought'' (Fodor, 1975) for sequences, akin to a mini computer language with primitives and rules whose combination can express any sequence of locations on an octagon (see Figure 1). The central idea is that the successive locations are not just encoded by their spatial coordinates, independently of one another. Rather, geometrical primitives of rotation and symmetry encode the transitions between items (Figure 1C), and an operator of repetition, akin to the ''for'' loop in programming languages, repeats the operations a certain number of times, possibly with variations (Table S1). As the language allows these instructions to be nested, it can express multiple levels of embedded repetitions and represent concepts such as ''square,'' ''rectangle,'' and ''two squares'' through a combination of numerical and geometrical information.

Our hypothesis is that whenever humans perceive a spatial sequence, they attempt to represent it in memory by searching for the simplest mental program that can generate it. Thus, the difficulty of memorizing a sequence should be proportional not to its length but to the length of its shortest generative program. The underlying hypothesis, which has been previously proposed and tested in many other contexts (Chater and Vitányi, 2003;

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

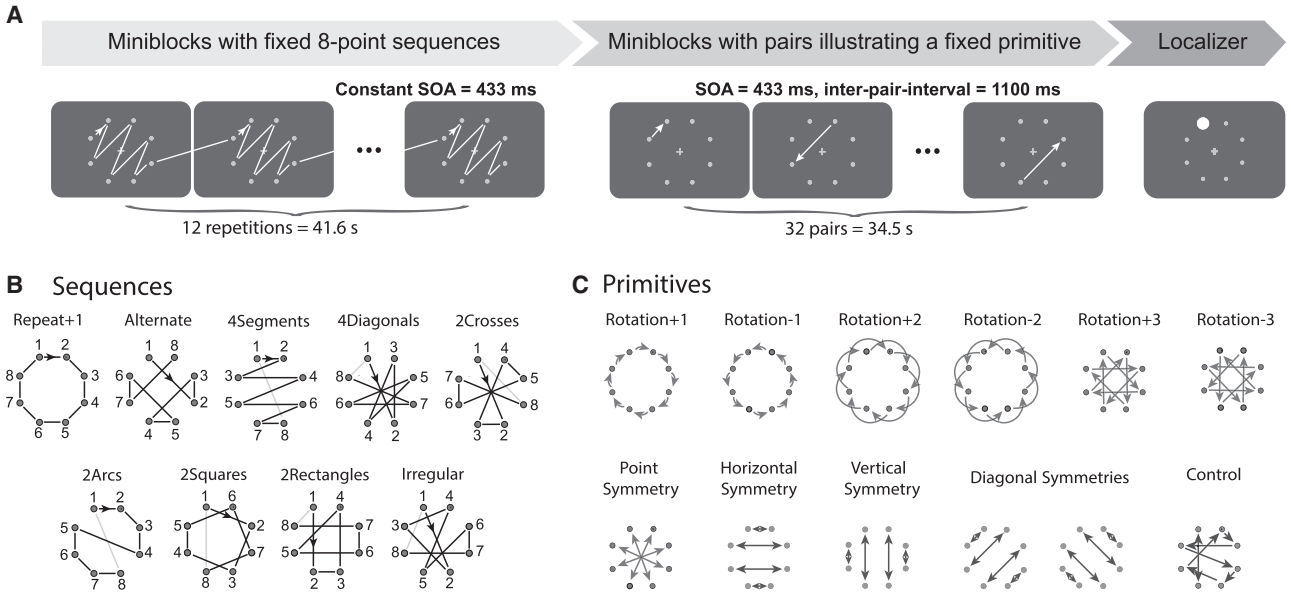

Figure 1. Experimental paradigm and stimuli

(A) The experiment was divided into three parts: (1) in the sequence part, a fixed spatial sequence of eight locations was repeated 12 times in each mini-block, and (2) in the primitive part, 32 pairs of two successive locations illustrating a given primitive were presented in each mini-block; participants were asked to report when they had identified the sequence or rule governing the pairs and to click whenever they detected a violation. Finally, (3) in the localizer block, dots were flashed at random locations on the octagon. The data were used to train a location decoder.

(B) Nine eight-location sequence templates were used in the sequence part. Presentation order is indicated by arrows. Actual sequences were generated by varying the starting point, rotation direction, and/or symmetry axis.

(C) The pairs of locations illustrating each of the 11 primitive rules presented in the primitive part. Arrows indicate the first and the second element of each pair.

Feldman, 2000; Leeuwenberg, 1969; Li and Vitányi, 1993; Mathy and Feldman, 2012; Romano et al., 2013), is that the brain operates as a "compressor" of incoming information that searches for the minimal description of incoming stimuli and that minimal description length is therefore a good predictor of psychological complexity. Hereafter, we refer to minimal description length of a sequence as the "language-of-thought complexity" or LoT complexity.

In a first behavioral study, we tested the hypothesis that participants encode spatial sequences using the proposed language of geometry (Amalric et al., 2017). Participants saw the beginning of a spatial sequence on the vertices of a regular octagon and had to predict the next locations. The results indicated that LoT complexity was a good predictor of behavior. Specifically, the more compressible a sequence was, the more subjects were able to anticipate on the upcoming item, even in the first trial, when they had never seen the entire sequence. Furthermore, error rate increased linearly as a function of LoT complexity, and the error patterns were compatible with the nested structure of the expression postulated by the formal language. Another group (Yildirim and Jacobs, 2015) also showed how a similar compositional language for spatial sequences could account for the transfer of abstract sequence knowledge from the visual to the auditory modality.

In a follow-up experiment (Wang et al., 2019), we measured fMRI activity while participants followed the same eight-location sequences with their gaze. Activity in the dorsal part of inferior prefrontal cortex correlated with the amount of compression, while the right dorsolateral prefrontal cortex (dlPFC) encoded the presence of embedded structures. Those brain regions belonged to a network involved in mathematical thinking and that is distinct from, but close to, the areas involved in natural language processing (Amalric and Dehaene, 2017). Although the content of sequences could not be decoded from fMRI signals, multivariate pattern analyses provided indirect evidence that the activity patterns in dorsal prefrontal cortex became increasingly differentiated as the sequences were learned.

fMRI is notoriously insensitive to the fine-grained timing of neural activity and thus failed to directly probe the precise temporal unfolding of language-like rules our theory predicted. In the present study, we therefore probed the existence of an abstract, language-like representation of geometrical sequences using magneto-encephalography (MEG), a sensitive technique with high temporal resolution. We used time-resolved multivariate decoding (King and Dehaene, 2014) and representation similarity analysis (RSA) (Kriegeskorte et al., 2008) in order to examine if the postulated geometrical and numerical primitives could be decoded at the precise moment when the postulated language of geometry suggests they should be deployed. Thus, we exposed human participants to several repetitions of geometrical sequences of variable LoT complexity. To ensure that they memorized the sequence, we asked them to click as soon as they had identified the repeating sequence and to detect occasional sequence violations. We then tested if MEG signals were sensitive to the postulated numerical and geometrical primitives. We identified markers of an anticipated representation of

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

CellPress

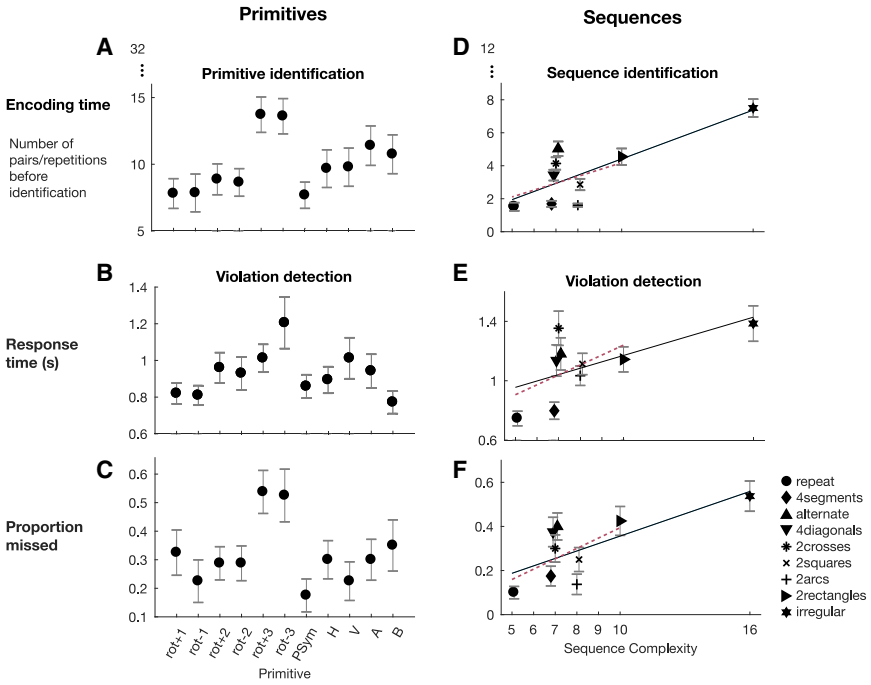

Figure 2. Behavioral performance and impact of sequence complexity

(A–F) For primitives and sequences, graphs show (1) the encoding time (i.e., the mean number of repetitions participants had seen before they identified the rule or the sequence) (A and D) and (2) the performance in violation detection as the mean response time to violations (top, B and E) and the proportion of missed trials (bottom, C and F). Error bars indicate the standard error of the mean (SEM). For geometrical sequences, linear regression lines indicate the effect of theoretical sequence complexity. The red dotted lines were obtained excluding the irregular sequence from the regression.

the sequence items and assessed their modulation by sequence LoT complexity. Moreover, we exposed participants to multiple exemplars of each primitive operation in isolation, and we probed if those primitives could be extracted from the MEG signals. This approach allowed us to determine if the sequences were indeed represented as nested repetitions of primitive rules.

RESULTS

Primitive rules: Behavioral results

During the primitive part of the experiment, participants were exposed to a succession of two dots forming a pair illustrating a primitive operation. For instance, in a given block, for all pairs, the second dot was vertically symmetrical to the first one, thus testing the primitive of vertical symmetry. All 11 primitives of the language were tested (see the list in Figure 1C; STAR Methods). We asked participants to click a button whenever they felt that they could predict the location of the second item in each pair. Those responses were converted into a measure of "encoding time" (Figure 2A): the number of pairs presented before the response occurred (if participants failed to respond, the maximum number of presentations, 32, was used). To assess if participants had understood the primitive rule and did not simply memorize the eight pairs by rote, we introduced a control condition in which the pairs were not driven by any general rule, and participants therefore had to memorize each of them. Participants performed very poorly in this condition, most of them failing to respond before the end of the run (i.e., after 32 pairs were presented). t tests showed that, for all of the 11 proposed primitives, encoding time was shorter than for the control condition (all p values < 10^{-5}), indicating that subjects detected all regularities.

In the second half of each block, we introduced a violation detection task: participants were asked to press another button as fast as possible whenever they detected that the second dot of a pair was misplaced (Figures 2B and 2C). In the control condition, participants missed 95% of those violations, whereas the average miss rate never exceeded 53% for the 11 geometrical primitives. Again, all the differences relative to control were significant (all p values < 10^{-4}).

To determine if all rules were processed with the same ease, we ran repeated-measures ANOVA on three dependent behavioral measures: encoding time, violation detection time, and violation miss rate. This analysis revealed significant differences among the primitives (respectively, $F[10] = 5.26$, $p < 10^{-4}$; $F[10] = 3.31$, $p = 0.001$; and $F[10] = 4.00$, $p < 10^{-4}$). Tukey post hoc tests on encoding time and miss rate indicated that the primitives of rotation ±3 were significantly more difficult than the counterclockwise rotation −1, the point symmetry, and the vertical symmetry (see Figure 2).

In summary, behavioral measures indicated that the participants could detect all of the postulated geometrical primitives but that, contrary to our initial assumptions (Amalric et al., 2017), those primitives may not be strictly equivalent in complexity, with rotation ±3 being more difficult to detect (for a similar conclusion, see Romano et al., 2018).

Geometrical sequences: Behavioral results

During the geometrical sequence block, subjects were repeatedly exposed to eight-item sequences. Subjects were asked to press a button when they had identified the repeating sequence precisely enough to predict the next item. We analyzed the encoding time, defined as the number of sequence repetitions that the participants needed before responding (if participants failed to respond, the maximum number of repetitions, 12, was used). As predicted, encoding time increased with LoT complexity, that is, minimal description length in the proposed language (Spearman $\rho = 0.45$, $t[19] = 9.6$, $p < 10^{-7}$; Pearson $r = 0.70$, $t[19] = 20.7$, $p < 10^{-13}$; Figure 2D; the results remained significant even when the most irregular sequence was excluded: $\rho = 0.22$, $t[19] = 3.1$, $p = 0.006$; $r = 0.30$, $t[19] = 4.3$, $p < 10^{-3}$) (Figure 2D).

After ten repetitions of a given sequence (i.e., during the last two presentations), a violation could occur (i.e., a single dot

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

was misplaced, off by three locations on the octagon). Participants were asked to press a button as fast as possible when they detected it. The violation detection times again exhibited a positive correlation with LoT complexity (Figure 2E; $\rho = 0.38$, $t[19] = 7.1$, $p < 10^{-5}$; $r = 0.40$, $t[19] = 6.7$, $p < 10^{-5}$; when excluding the irregular sequence: $\rho = 0.33$, $t[19] = 6.2$, $p < 10^{-5}$; $r = 0.28$, $t[19] = 5.2$, $p < 10^{-4}$). The effect was also detectable in error rates: as the sequence LoT complexity increased, so did the number of missed violations (Figure 2; $\rho = 0.33$, $t[19] = 6.3$, $p < 10^{-5}$; $r = 0.43$, $t[19] = 7.1$, $p < 10^{-6}$; excluding the irregular sequence: $\rho = 0.18$, $t[19] = 2.4$, $p = 0.03$; $r = 0.28$, $t[19] = 3.8$, $p = 0.001$). Indeed, the effect on errors was very large: subjects missed fewer than 10% of deviants in the simplest sequence but more than 50% of them in the most complex sequence, consistent with the idea that the latter exceeded their working-memory span and they had trouble memorizing it.

As all sequences were of the same length, those results cannot be explained by classical slot-based models of working memory. They converge with the ones obtained using explicit predictions of the next item (Amalric et al., 2017) or using eye-movement anticipations (Wang et al., 2019): the more complex a sequence, the harder it is to predict the next locations and, therefore, to detect violations. Nevertheless, the correlations with LoT complexity were modest, suggesting that our measure of LoT complexity, based on the proposed language, may not be ideal. Indeed, the behavioral results on the primitive part indicated that rotation ±3 was harder to process than the other primitives. Furthermore, participant's behavior may also be modulated by other parameters such as the spatial distance between items. To determine the contribution of these two variables to behavioral performance, we ran a stepwise regression (STAR Methods) to assess which linear model minimized the AIC (Akaike information criterion). The best model of encoding times and response times was one that included both LoT complexity and presence of rotation ±3. The best model for the proportion of missed violations contained all three predictors: LoT complexity, presence of rotation ±3, and distance. This observation was confirmed when we computed a more empirically driven measure of complexity from the intercorrelation between our three dependent measures. Both detection time and missed rate were well predicted by encoding time, with correlations similar to those obtained from our theory-driven measure of LoT complexity (respectively, $\rho = 0.54$, $t[19] = 10.5$, $p < 10^{-8}$, $r = 0.51$, $t[19] = 9.8$, $p < 10^{-8}$, and $\rho = 0.44$, $t[19] = 5.7$, $p < 10^{-4}$, $r = 0.48$, $t[19] = 5.4$, $p < 10^{-4}$). We thus extracted the first principal component of those three dependent behavioral measures (Figures 2D–2F). Although determined in a purely data-driven manner, this empirical measure of complexity showed a robust correlation with the theory-driven LoT complexity ($\rho = 0.56$, $r = 0.75$). Deviations from a perfect line were due to deviations for the alternate, four-diagonals and two-crosses sequences, which either contained the rotation ±3 primitive or long spatial distances. To conclude, the proposed language should be slightly amended to allow for complexity that varies with distance and primitive type.

Decoding the successive locations of each sequence item

We first tested whether MEG signals contained decodable information about the successive locations of each sequence item. At each time point, on the basis of the 306 sensor measures, we trained an estimator to decode the angular position of the presented item. This position decoder was trained on data independent of the eight-item sequences and for which no anticipation could be formed (STAR Methods).

As shown in Figure 3A, position decoding was at chance prior to stimulus presentation, rose suddenly ~70 ms after stimulus onset, and peaked at 150 ms. Successful decoding of the eight item locations was based primarily on bilateral occipito-parietal sensors (Figure S7). The generalization-across-time (GAT) matrix (Figure 3B) revealed both a diagonal, indicating an unfolding sequence of stages, and a partial square pattern, indicating a sustained maintenance of location information in brain signals (King and Dehaene, 2014). Successful generalization to the eight-item sequence data was also observed (Figure 3C). We trained a decoder on the average brain responses in the time window of maximal decodability, from 100 to 200 ms after the stimulus onset, and tested it on each sequence. Figure 3D presents the relative amount of times that the decoder predicted each location for each ordinal position. The pattern obtained from these predictions tightly paralleled the actual profile (Figure 3D, second line), with only some added spatial uncertainty (i.e., spreading of the decoding to the two neighboring locations on the octagon). Control analyses showed that this decoding did not arise from residual eye movements (Figure S1).

Prior research on predictive coding has demonstrated that predictable stimuli elicit a reduced brain response but a more faithful representation, as reflected by a higher decoding accuracy (Kok et al., 2012; Summerfield and de Lange, 2014). To test this prediction, we examined if simpler sequences elicited higher decoding accuracy. We ran a linear regression of the average decoding score as a function of LoT complexity and empirical complexity for each participant. A small modulation was found as a function of empirical complexity (one-tailed t test, $t[19] = -2.0$, $p = 0.0297$) but not of LoT complexity (one-tailed t test, $p = 0.1$, $t[19] = -1.3$).

Anticipation and its modulation by sequence structure

Our next goal was to determine if anticipatory information was present on MEG even prior to actual stimulus presentation, as previously demonstrated by others (Demarchi et al., 2019; Ekman et al., 2017; Kok et al., 2014, 2017) and if it decreased with sequence LoT complexity.

To do so, we assessed the performance of the position decoder prior to the presentation of each sequence item (time 0 ms). Importantly, as there was evidence of spillover of the decoding to nearby locations on the octagon, we controlled for distance to the previous item. To this aim, we defined an anticipation score as the difference in the decoding score at the upcoming location and at the location equidistant from the previous location but that was not stimulated (STAR Methods; Figure 4A). This anticipation score was computed for all sequence items (Figure 4B) then averaged across the training time window 100–200 ms (Figure 4C), corresponding to the maximal performance of the position decoder (Figure 3A). We ran a cluster-based permutation test in the temporal window between the presentation of the preceding item and the anticipated item (i.e., from −430 to 0 ms). The anticipation score was significantly

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

CellPress

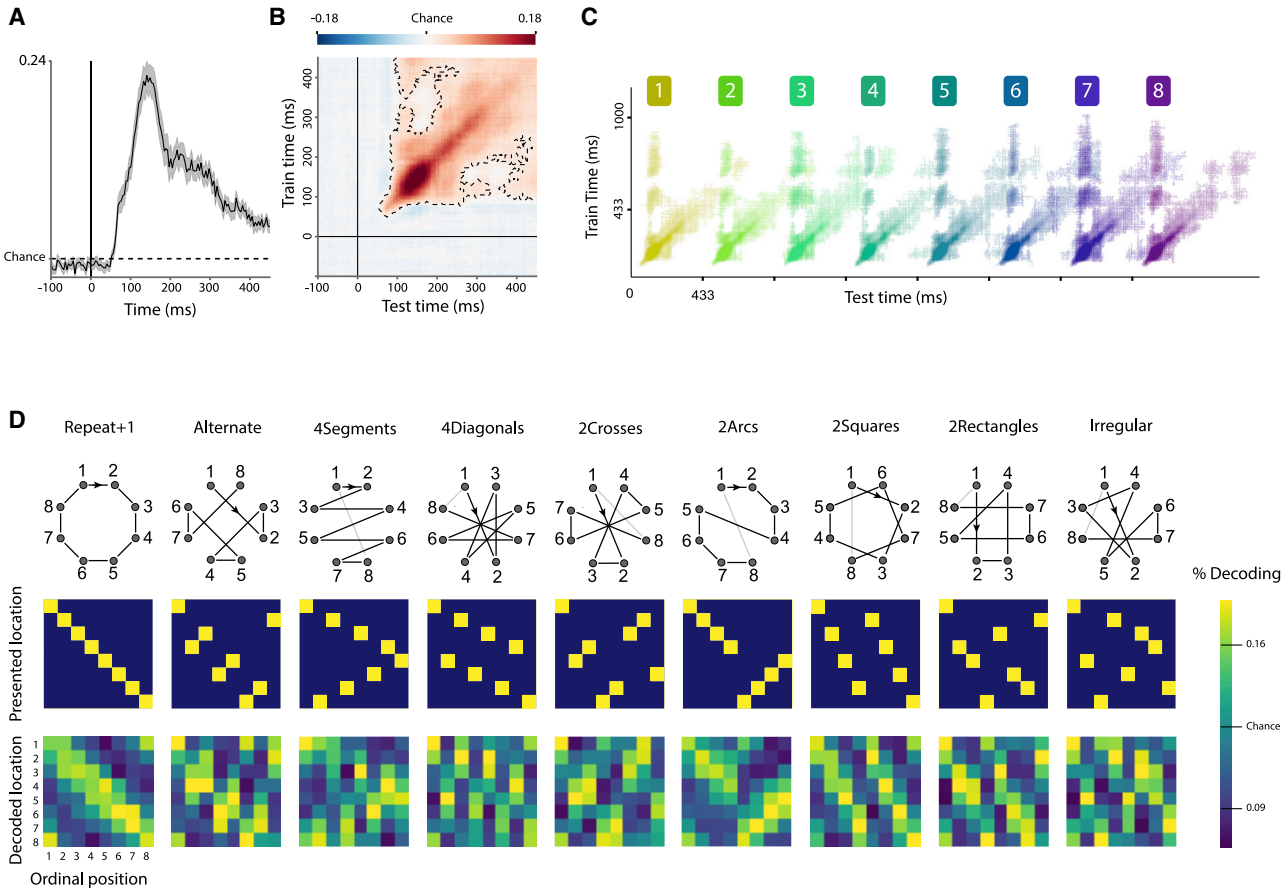

Figure 3. Decoding successive sequence locations from MEG signals

(A) Performance in decoding stimulus location as a function of time following the flash of a dot at a given location (the shaded area indicates the sem). Maximal decoding performance was achieved at ~150 ms.

(B) Average generalization-across-time (GAT) matrix showing the location decoding score as a function of training time (y axis) and testing time (x axis). The dashed lines indicate $p < 0.05$ cluster-level significance, corrected for multiple comparisons (STAR Methods).

(C) Thresholded ($p < 0.05$, corrected) decoding matrix plot showing when each of the eight successive sequence items could be decoded.

(D) Decoding of each sequence. The top matrix represents the stimuli at each ordinal position, and the bottom matrix shows the proportion of times a given spatial location was decoded.

positive during two time windows, from −176 to −104 ms and from −84 to 0 ms. Additional analyses excluded a contribution of eye movements to those anticipation signals (Figure S1).

We then used anticipation to probe the representation postulated by the language of geometry. If predictive mechanisms are modulated by sequence structure, the anticipation score should be increasingly smaller as the sequence gets more complex. We computed the average anticipation score for each sequence (except four diagonals and two crosses; STAR Methods) and ran a linear regression as a function of complexity. We observed a significant decrease of anticipation score with LoT complexity (one-tailed t test, $t[19] = -2.1$, $p = 0.025$) and empirical complexity (one-tailed t test, $t[19] = -2.0$, $p = 0.028$). This result, although only weakly significant, supports the hierarchical representation postulated by the language of geometry, as it shows that expectation mechanisms, measured by the anticipation score, are modulated by the overall LoT complexity of the sequence.

Previous research has shown that brain activity in language areas is modulated by the nested structure of language, such that activity varies sharply at the boundary of sentence constituents such as noun phrases (Nelson et al., 2017). We thus wondered if a similar effect occurred with the language of geometry. To do so, we therefore compared the anticipation scores of the items that, according to our postulated language, open a constituent with the ones that are inside a component. Only a small amount of data conformed to those conditions, however (STAR Methods), and perhaps because of this lack of statistical power, no effect of syntactic structure reached significance.

Decoding geometrical operations

The language-of-geometry hypothesis predicts that participants encode spatial sequences not only in terms of each item's specific location but also in terms of high-level geometrical primitives such as symmetries and rotations. To directly probe the existence of this representation, we attempted to decode,

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

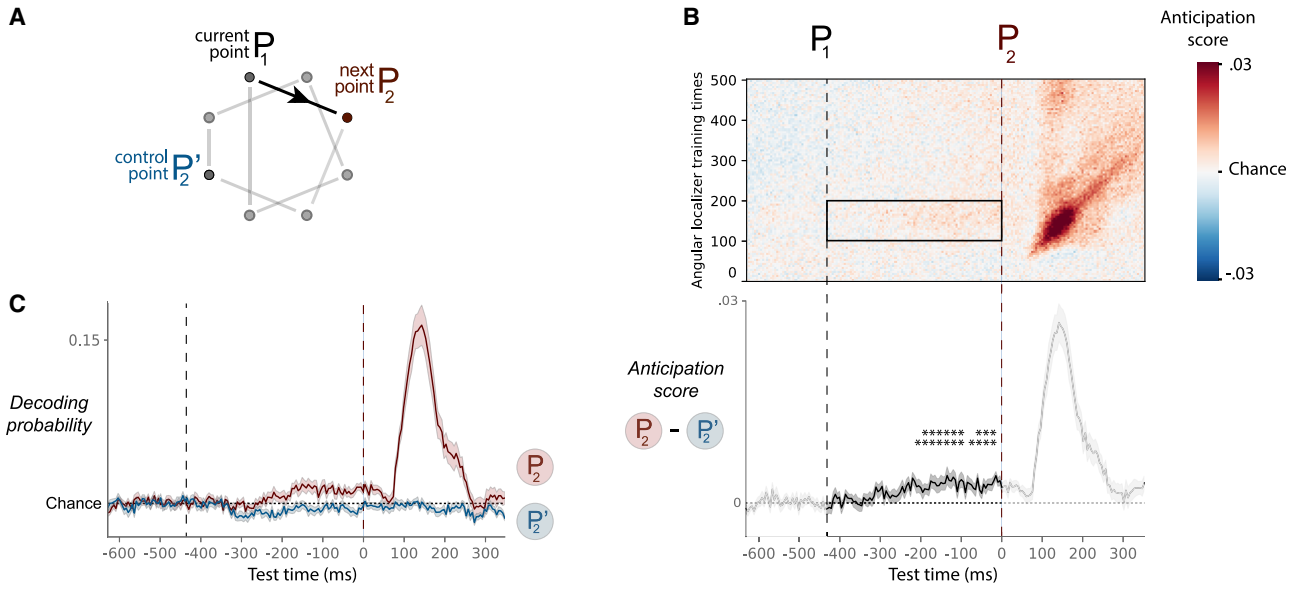

Figure 4. Detecting an anticipation of sequence locations from MEG signals

For each training time in the range 0–500 ms, we tested whether the location decoder could detect the stimulus location in a time window ranging from 630 ms before to 370 ms after stimulus presentation.

(A) To control for distance from the previous stimulus location (P_1), an anticipation score was computed by measuring the decoding at the correct stimulus location (P_2) and subtracting the decoding at the equidistant non-stimulated location (P'_2).

(B) Top panel: the anticipation score is above chance before stimulus presentation. Bottom panel: average anticipation score over 100–200 ms training times (selected region from top panel). The anticipation score was significantly above chance before stimulus presentation (time window, asterisks), indicating that the brain anticipates P_2 even before it appears (shaded area indicates the sem).

(C) Average of the proportion of times P_2 and P'_2 were predicted over the 100–200 ms training time window. The average anticipation score is the subtraction of the two (shaded areas indicate sem).

from the brain signals, the elementary geometrical primitives postulated by our formal language.

To do so, we first examined if the 11 primitive operations could be decoded when they were presented in isolation. Using the trials that illustrated each primitive (see Figure 1C), we trained a decoder to determine, on the basis of the brain signals, which of the 11 primitive operations was presented (STAR Methods). Our hypothesis was that participants actively apply the appropriate geometrical transformation to predict the location of the second element. To characterize the temporal dynamics of the primitive operation code, the primitive decoder was trained and tested at different time points. We cross-validated across runs to exclude decoder overfitting due to temporal proximity, and we ran a sliding window over the epochs to increase the signal-to-noise ratio (STAR Methods). As shown in Figure 5A, performance was significantly above chance ($p < 0.05$ cluster-level significance) for an extended time window, which peaked ~200–300 ms following the presentation of the first element of the pair but actually started ~50 ms before that presentation, suggesting that the block structure enabled participants to anticipate on the forthcoming geometrical transformation. Thus, those results indicate that human brain activity contains decodable information about the type of geometrical transformation that links one location to the next and does so, in a predictable context, even before the second sequence item is presented.

The 11 geometrical primitives in Figure 1C are all perfectly balanced in terms of starting point and endpoint, and their decoding is therefore unconfounded by retinotopic stimulation. However, by definition, they involve different pairs of locations, and some primitives differ in the distance between the two locations (e.g., +1, +2, +3). It is unclear whether this represents a genuine confound, because the time window when geometrical transformations were decoded preceded the presentation of the second item of the pair and hence came before distance could have any physical effect (e.g., perceived motion). Still, we wondered whether, in the extreme case, geometrical transformations could be decoded even if the pair of starting points and endpoints were strictly identical. We capitalized on the fact that the same pairs of dots could appear in the context of either a rotation block or a symmetry block. For instance, a dot moving from the top left to the top right location can be construed as either a rotation around the octagon (+1) or as a symmetry with respect to the vertical axis (see Figure 5C). We selected all trials corresponding to such pairs and asked whether, for equal start and end locations, brain activity still contained decodable information about their putative internal encoding as a rotation or as a symmetry (STAR Methods; Figure 5C). Figure 5D shows that this was indeed the case. These decoding results suggest that over and above any location or distance information, the neural patterns associated with abstract geometrical operations of rotation and symmetry can be disentangled when primitive operations are considered in isolation.

We then determined whether, on the basis of the description provided by the postulated formal language, the same primitive

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

CellPress

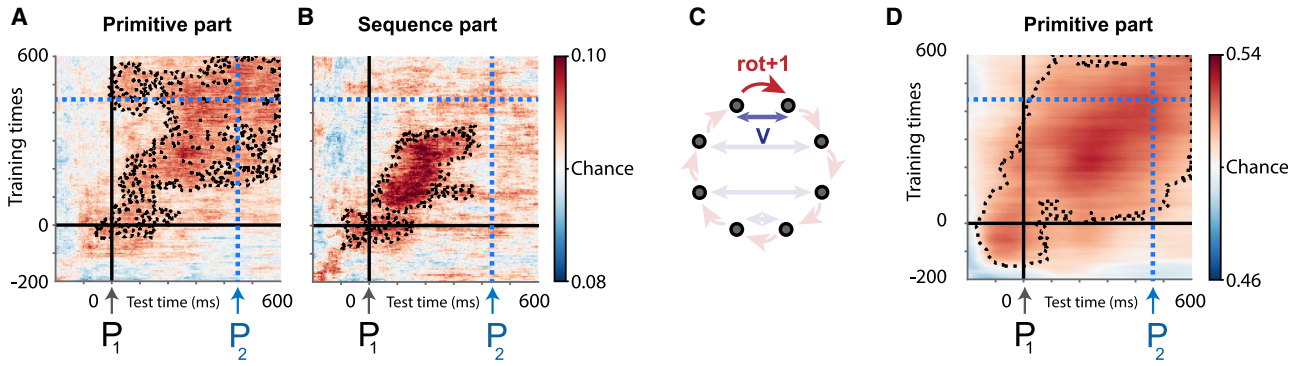

Figure 5. Decoding primitives

Decoders were trained to predict which of the postulated geometrical transformations was applied at a given time between two consecutive locations P_1 and P_2 (onset of first location marked by t = 0 ms).

(A and B) Average GAT matrices showing the decoding score for the 11 primitive operations in the primitive part of the experiment (A) and in the sequence part (B).

(C) Illustration of how trials with the same start and end locations may be classified as a rotation or a symmetry depending on the context.

(D) Performance of a binary decoder for rotation versus symmetry on such trials with identical start and end locations, from the primitive part of the experiment. Dashed lines indicate $p < 0.05$ cluster-level significance over the −200 to 600 ms time window, corrected for multiple comparisons (STAR Methods).

operations could be decoded in the context of a sequence. We extracted 800 ms epochs centered on specific locations in the sequence and labeled them by the subsequent primitive operation involved, at the lowest level of the sequence description (i.e., the one whose application predicted the next location). We successfully trained a decoder for the 11 primitives, as indicated by an above-chance GAT performance matrix (Figure 5B). However, when we replicated the above training and testing on a balanced set of rotation and symmetry pairs, we could not detect, at above-chance level, a code separating rotations from symmetries, perhaps because of the smaller number of trials involved.

We then assessed whether the primitive code, inferred when such primitives were presented in isolation, generalized to the sequence part of the experiment, when elementary primitive operations are integrated into a program. In order to avoid any bias due to a particular block or run, we trained this decoder on the micro-averaged trials over the four runs (STAR Methods) and tested it on the sequence data. The GAT matrices did not exhibit any significant clusters, suggesting that the neural code for primitives presented in isolation is not directly replicated in the context of a full sequence but is modified in the sequence context (see Discussion).

In summary, the MEG decoding analysis provided evidence for an abstract encoding of rotations and symmetries, independently of the visual features of the stimuli, both in the primitive and in the sequence parts of the experiment. However, this code did not generalize from the primitive to the sequence part.

We also used RSA (Kriegeskorte et al., 2008) to further test the existence of a neural code for abstract geometrical primitives, independently of a location-specific code. To do so, we selected epochs corresponding to transitions encoded by an unambiguous primitive according to our language. The representational dissimilarity matrix, computed separately for each primitive operation and each location on screen, was regressed as a function of several theoretical predictors (see Figure S2; STAR Methods). A cluster-based permutation test on the 0–1 s time window confirmed that, over and above retinotopic or visuospatial factors (the locations of the first and the second item of the pair and the distance between them), a representation of abstract geometrical primitives influenced brain activity. Regression coefficients for the primitive operation were significantly positive in the window between 100 and 400 ms (i.e., between the presentation of the two items of the pair), during which an application of the corresponding transformation allowed to predict the location of the second item. This representation was reactivated after the second item was presented (significant cluster from 710 to 870 ms). We ran a cluster-based permutation test from −400 to 0 ms and found a significant cluster from −400 to −250 ms, suggesting that the primitive operation was represented before the onset of the first item of the pair. We then used RSA to again ask whether the neural code for primitive operation was similar across experimental parts. Although a mild peak was observed, it did not reach significance after correction for the time interval tested. Thus, consistent with the decoding results, the neural representation of geometrical primitives differed in the context of a sequence and when presented in isolation.

Decoding ordinal position in subsequences

According to the proposed hypothesis, the internal representation of the sequence is not a mere list of locations but expresses nested structures. For instance, "two squares" is encoded as "two groups of four items, each linked by a +2 operation." When remembering such a sequence, an internal numerical code, akin to the for loop in programming languages, must unfold in the participants' brain, keeping track of how many times a given geometrical transformation has been applied.

Thus, we investigated the presence, in human MEG signals, of a neural code for ordinal position within a subsequence of items, ranging from one to a maximum of four. The presence of such a code in MEG signals would reflect the parsing of the sequence into multiple subgroups linked by a common geometrical transformation, similar to the parsing of language sequences into

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

nested phrases (Ding et al., 2016). An ordinal number code has been previously observed in a simpler context in both human and non-human primates (Kutter et al., 2018; Nieder, 2012; Nieder et al., 2006) and has been postulated in some models of working memory (Botvinick and Plaut, 2006; Botvinick and Watanabe, 2007).

Our analyses focused on the two-squares and two-arcs sequences (two groups of four items) and on the four-segments and four-diagonals sequences (four groups of two items). In each case, we trained and tested a decoder on sequence data labeled by the innermost ordinal position (STAR Methods). The GAT matrices, obtained by cross-validating across runs, showed significant above-chance decoding (Figures 6A and 6C, left column; corrected $p < 0.05$ with cluster-level permutation analysis), suggesting that ordinal position was indeed encoded in MEG signals. Performance was above chance before the stimulus was actually presented, compatible with the fact that sequence items were anticipated.

We then determined if the code for ordinal position is abstract and thus identical regardless of the particular geometrical primitive that is called for at that position. In other words, we asked whether the code for a sequence is "factored out" into separate codes for ordinal position and for the particular geometrical primitive applied at this position, as predicted by our language model and as previously shown for non-spatial visual sequences by Liu et al. (2019). Indeed, we found above-chance generalization across time and across sequences (Figures 6B and 6D, left column), indicating that the code for ordinal position is, at least in part, independent of the specific geometrical transformation involved.

To characterize the temporal variations of this code during the entire sequence presentation, we applied the ordinal position decoders to MEG data from runs that were not used for training. As a more sensitive estimator of classifier performance, we computed the mean projection of the data on the decoding axis for each predicted ordinal position. We averaged the predicted distances on the 300–500 ms time window for which the decoding score was above chance. The results (middle columns) indeed showed oscillations compatible with a sequential unfolding of an ordinal position code across the entire sequence, for decoders trained on both sequences (Figures 6A and 6C) and when generalizing across sequences (Figures 6B and 6D). To test if the ordinal code indeed oscillated at the appropriate frequency, we computed the log power spectrum of the distance time series for each condition. We then determined if there was a significant peak at the component frequency (two arcs and two squares, f/4 = 0.58 Hz; four diagonals and four segments, f/2 = 1.15 Hz) by comparing the log power at that frequency with the neighboring frequencies (STAR Methods). In the case of two squares and two arcs, the test was performed on the averaged log power over the four ordinal position decoders. In every condition, the test was significant at the predicted frequency, indicating the presence of groups of two or four items depending on the sequence (four groups of two items: [A] cross-validation across blocks $t_{f/4}[19] = 0.70$, n.s., $t_{f/2}[19] = 3.38$, $p_{f/2} < 0.01$; [B] generalizing across sequences $t_{f/4}[19] = -0.59$, n.s., $t_{f/2}[19] = 3.32$, $p_{f/2} < 0.01$; two groups of four items: [C] cross-validation across blocks $t_{f/4}[19] = 6.27$, $p_{f/4} < 0.001$, $t_{f/2}[19] = 2.28$, $p_{f/2} = 0.034$; [D] generalizing across sequences $t_{f/4}[19] = 3.91$, $p_{f/4} = 0.001$, $t_{f/2}[19] = 1.35$, n.s.).

Furthermore, a 2 × 2 ANOVA of the log power difference, with factors of frequency (f/2 or f/4) and sequence type (components of size two versus four), showed a significant interaction both when training on both sequences ($F[1, 19] = 20.0$, $p = 2.6 \times 10^{-4}$) and when generalizing across sequences ($F[1, 19] = 14.7$, $p = 0.001$), indicating that the power was significantly stronger at the expected than at the inappropriate frequency (STAR Methods).

Instead of ordinal number, the brain might have solely encoded the position of the first and the last item in each group. However, we could reject this hypothesis because significant oscillations were also present, with a peak at the appropriate time, for the intermediate positions 2 and 3 in four-segment and four-diagonal sequences (Figure S3). Nevertheless, examination of the cross-generalization between the ordinal codes for groups of two versus four items suggested that first versus last item information was also encoded (Figure S4).

As a final control, we trained and tested the very same decoders on data for which our language model predicts that there should be no subgroups of items (i.e., the repeat and irregular sequences). As predicted, the decoders now failed to identify a 1-2-1-2 structure (Figure S5). Interestingly, there was modest evidence for a 1-2-3-4 code only in the repeat +1 sequence, suggesting that even when the items keep going around the octagon without any break, participants may encode them in memory as groups of four (e.g., top versus bottom or left versus right). Crucially, the decoders found no evidence of ordinal coding in the irregular sequence.

Finally, Figure S7 shows that the spatial, primitive, and ordinal decoders relied on partially distinct MEG sensors. While occipital-parietal sensors, located over posterior retinotopic maps, contributed in a preponderant manner to the spatial code, ordinal codes were associated primarily with right precentral sensors. Primitive codes were less focal and involved distributed frontal, parietal, and temporal sensors.

DISCUSSION

The goal of the present study was to probe the internal representations that humans use to encode geometrical sequences of varying regularity. The simplest models of working memory for serial order assume either that sequences are encoded as simple associative chains linking consecutive items (Lewandowsky and Murdock, 1989) or by storing each item in a distinct memory slot (Botvinick and Watanabe, 2007; for review, see Hurlstone et al., 2014). If this were the case, however, all of our sequences would be encoded in a similar manner, as they all have the same length and only differ in the order in which the same eight locations are presented. Instead, we found evidence that participants mentally compress the sequences using their geometrical regularity and have a better memory for those that can be compressed down to a lighter memory load. Indeed, both behavioral and brain-derived measures were modulated by LoT complexity, as provided by the postulated formal language. Furthermore, using time-resolved decoding and RSA techniques, we identified three distinct types of codes: for each item's spatial location,

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

CellPress

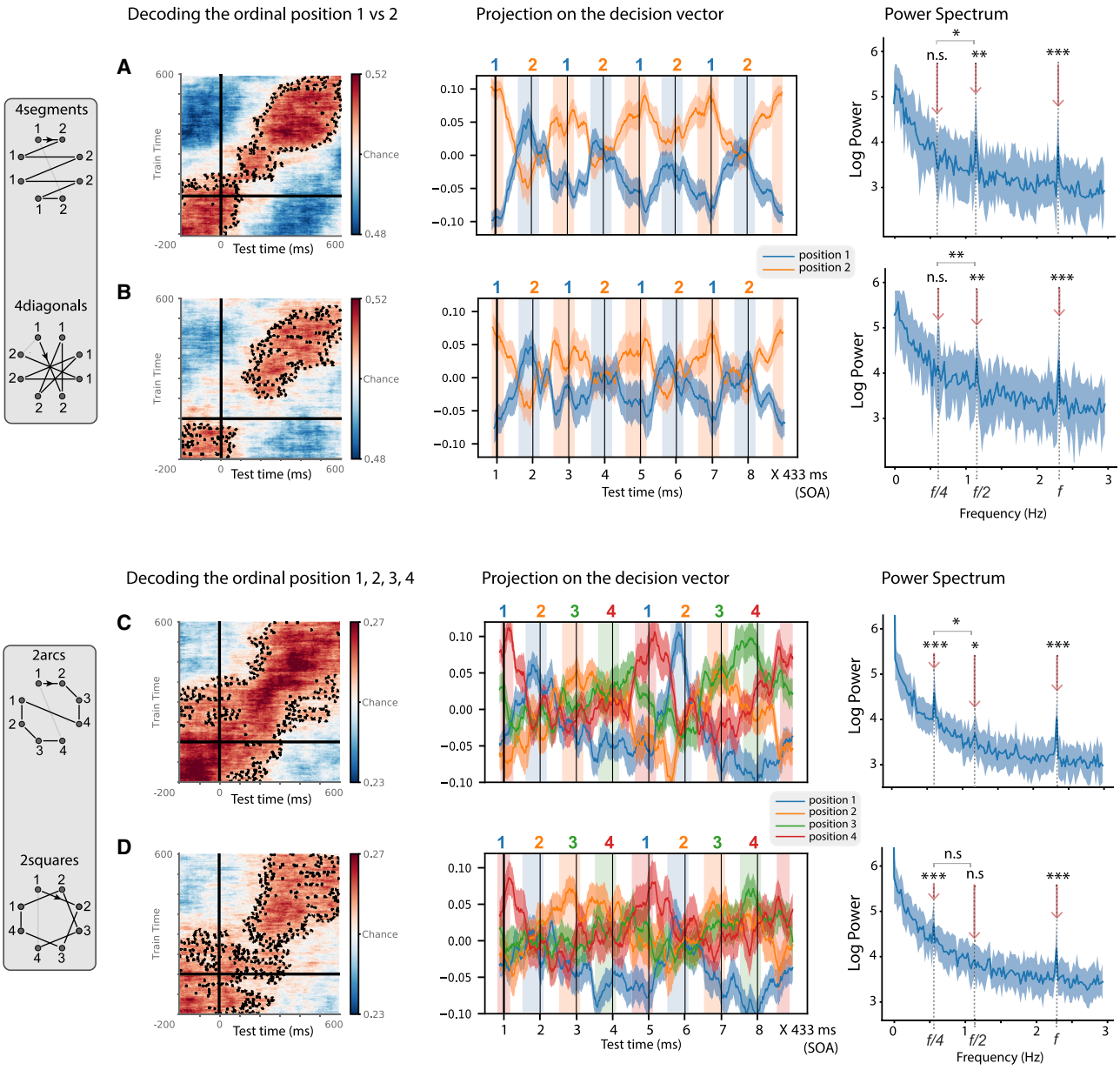

Figure 6. Decoding ordinal position within a sequence component

(A–D) Decoders were trained to predict the ordinal positions 1 and 2 in the inner groups of four-segment and four-diagonal sequences (A and B) and positions 1–4 in the inner groups of two-arc and two-square sequences (C and D). (A) and (C) were obtained by cross-validating across blocks and (B) and (D) by generalizing across sequences (e.g., from four segments to four diagonals and vice versa). Left column: mean GAT performance for the ordinal position decoder. Zero milliseconds corresponds to item onset. Dashed lines indicate $p < 0.05$ cluster-level significance, corrected for multiple comparisons (STAR Methods). Middle column: time course of the average output of ordinal decoders trained in a time window of 300–500 ms, during a sequence of eight consecutive items. Although noisy, those curves show a decoding peak around 300–500 ms after the corresponding ordinal item (colored zones) and a clear rhythmicity every two items for the top sequences (A and B) and every four items for the bottom sequences (C and D). The y axis shows the projection on the decision axis for each decoder, and the vertical lines indicate the onsets of sequence items (thick line marks the first item). Blue, orange, green, and red lines indicate first, second, third, and fourth ordinal positions. Shaded areas indicate the sem. Right column: power spectrum of those time courses throughout the presentation blocks for the two sequences. Statistics are provided at frequencies f = 2.31 Hz, f/2 = 1.15 Hz, and f/4 = 0.58 Hz, where f is the presentation frequency.

for the geometrical transformation linking two consecutive locations, and for their ordinal position within a group of items.

Using multivariate decoding, we found that each of the successive retinotopic locations in a spatial sequence could be decoded (Figure 3). However, we also found that the brain did not stop at encoding specific locations but also coded for the transitions between consecutive locations, in such a way that we could classify them into 11 abstract geometrical primitives

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

(Figure 5). Decoding of geometrical primitives worked both when considering primitive operations in isolation and in the context of a sequence. In addition to evidence from decoding, RSA allowed to factor out the contributions of retinotopic and abstract geometrical codes and to demonstrate that even if the first one dominated the MEG signal, the second, although weaker, was significant (Figure S2).

Our findings support the existence of a mental repertoire of abstract geometrical concepts and its automatic deployment when a sequence must be encoded in memory. A previous behavioral study showed that even young children and adults with limited access to formal schooling in mathematics use these geometrical primitives when memorizing sequences (Amalric et al., 2017). All humans, starting at an early age, may have access to a core of basic geometrical concepts (Dehaene et al., 2006; Spelke et al., 2010).

Another prediction of our formal language is that the human brain uses such geometrical regularities to segment the sequences: if a series of successive locations can be encoded by the repeated application of the same transformation, then subjects compress it using an internal repetition operator, akin to a for loop in programming languages. In agreement with this hypothesis, MEG signals contained decodable information about ordinal position within a subgroup of locations (Figure 6). The ordinal code was activated in a periodic manner (Figure 6, middle column) in agreement with the proposed rhythmicity of the sequence representation, which varied across sequences (groups of two or four). Such ordinal knowledge was not present in every sequence but could be decoded only from brain signals when the code was indeed predicted by our language model (compare Figure 6 and control in Figure S5).

Previous experiments have revealed that a numerical code is present in several cortical regions of the monkey (Nieder, 2012; Nieder et al., 2006; Ninokura et al., 2004) and human brain (Fias et al., 2007; Kutter et al., 2018; Nieder and Dehaene, 2009). Here, we also found suggestive evidence for the decoding of the first and the last ordinal positions. Those positions also correspond to the opening and closing of a component and may therefore reflect the distinct mental operations involved. This finding fits with the ubiquity of primacy and recency effects in working memory (Anderson et al., 1998; Hurlstone et al., 2014; Orlov et al., 2000; Terrace et al., 2003). The ordinal code for intermediate positions 2 and 3 could also be decoded and recurred in a periodic manner (Figure S3), indicating that the ordinal code was not limited to first versus last.

Although we refer to an ordinal code, we readily acknowledge that because the sequences unfolded at a fixed pace, our experiments could not separate numerical and temporal codes. Future experiments could use variable sequence length, tempo, and grouping in order to better specify the nature of the neural code involved. Importantly, all such codes point to the same conclusion, namely, that the human brain does not stick to a flat, superficial representation of the sequence but parses it into subsequences on the basis of geometrical cues. In this respect, our results extend the conclusions obtained for linguistic structures in connected speech (Ding et al., 2016) to the domain of visuospatial sequences.

Another important postulate of our proposed language of thought is that different sequences can be encoded via a recombination of the same ordinal and geometrical codes. In a recent MEG study, Liu et al. (2019) proposed the term "factorized codes" for such a situation in which a sequence is encoded as the combination of an abstract structural code (here for the ordinal number 1–4) and a specific content (here the specific geometrical transformation, e.g., +1 to generate an arc or +2 to draw a square). In their study, Liu et al. (2019) exposed participants to sequences of pictures and asked them to reorder them according to a learned rule. They showed that human MEG signals contain three different codes: for item identity, item ordinal position in the sequence, and target sequence after application of the requested mental transformation. The hippocampus replayed the learned structure, supporting the hypothesis of a factorized hippocampal representation of abstract structural sequence knowledge. The current results are consistent with Liu et al., (2019) findings, as we found that the structural code that tracks ordinal position generalizes across sequences that make use of different geometrical primitives, suggesting that ordinal knowledge is encoded in an abstract manner, independently from the code used to represent the particular geometrical primitive. This subdivision of labor fits with our formal language, in which the same for loop can be applied to different geometrical primitives.

A final prediction of our language is that those primitives can be nested to form complex, recursive mental programs and that working memory load is ultimately determined by the minimal description length of that program, a measure we call LoT complexity (Planton et al., 2021). We obtained behavioral and MEG evidence in support of this prediction. Behaviorally, participants' subjective feeling of remembering the sequence, as well as their objective capacity to detect an occasional location deviant, were correlated with LoT complexity (Figure 2). In MEG, we found that covert visual anticipation of the sequence items was modulated by sequence complexity, suggesting that expectation mechanisms can be more efficiently deployed when the sequence has a low LoT complexity. The observation of anticipation signals fits within the predictive-coding framework, which proposes that the brain constantly tries to predict its sensory inputs (Bastos et al., 2012; Chao et al., 2018; Friston et al., 2003) and projects stimulus-specific templates in advance of the stimulus itself, as previously found during associative learning (Demarchi et al., 2019; Kok et al., 2014, 2017; Sakai and Miyashita, 1991) or during sequences of spatial locations with no structure (Ekman et al., 2017). Our previous behavioral studies showed that human participants could anticipate upcoming geometrical sequence items, both explicitly (by pointing) and implicitly (by moving their gaze), and that such anticipation varied according to the hierarchical structure of the sequences (Amalric et al., 2017; Wang et al., 2019). Wang et al. (2019) further showed that fMRI signals in bilateral inferior frontal gyrus (IFG) and dlPFC were modulated by LoT complexity and by the actual amount of anticipation of the nested rules. These regions were previously shown to be engaged in the parsing of structured spatial sequences (Bor et al., 2003; Desrochers et al., 2015), but those studies considered only two levels of sequence regularity (structured versus unstructured), as opposed to the

multiple levels of regularity and nesting studied here. The results from Wang et al. (2019) bolster the proposition of a hierarchical caudal-rostral organization of the prefrontal cortex to represent rules of increasing abstraction (Badre, 2008; Badre and D'Esposito, 2009; Badre and Nee, 2018; Badre et al., 2010; Koechlin and Jubault, 2006; Koechlin et al., 2003). These prefrontal regions may encode the internal model of geometrical sequences and send top-down signals in order to pre-activate circuits in posterior cortices, generating the observed anticipations.

Several limits of the present work must be acknowledged. First, the formal language of geometry we proposed provided only an imperfect, though significant, fit to the observed behavioral data (Figure 2). This deviation of behavior from theory seems to arise, at least partially, from the fact that some primitives, such as rotation ±3, are more difficult to process than others. In an augmented version of the theory, weights could be assigned to each primitive separately before computing sequence complexity (as proposed by Romano et al., 2018).

A second limit is that we did not provide direct decoding evidence that several numerical and geometrical primitives could be jointly encoded in a complex program. So far, we managed to decode only the lowest level of this hypothetical nested code (i.e., the transitions between consecutive items and their ordinal positions within a local group of locations). Whether and how the brain binds the same representation at multiple levels to create a nested language-like code is a debated issue (e.g., Elman, 1990; Marcus et al., 1999; Smolensky and Legendre, 2006) that remains to be empirically resolved. In this context, it is noteworthy that the decoder of geometrical primitives, once trained on the data from the primitive part, did not generalize to the sequence part where the same primitives were embedded into longer sequences. This result was confirmed by RSA. Future research should disentangle two alternative explanations. First, this negative finding could be due to peripheral factors such as the different timing of the two parts (Figure 1) or the possibility of head position changes across the experiment. Alternatively, it may also hint at a principled difference in the neural encoding of the same geometrical primitives when presented in isolation and in a sequence context. Indeed, a major difference is that in the primitive part, a single primitive rule is considered at the time, for an entire block, whereas in the sequence part, the primitives must be bound together with other primitives and with numerical information to form complex expressions such as "two squares." Some theories postulate that to embed an object inside a syntactic structure, the brain applies a tensor product operation between the neural codes for the object and for its role inside the structure (Smolensky, 1990; Smolensky and Legendre, 2006). This radical transformation would re-map the original neural vectors for primitives onto a new direction in neural space, thus explaining why they can no longer be decoded by the original decoder once they are embedded in a sequence. Higher resolution recordings, possibly at the single-cell level, may be required to test this hypothesis.

The present work fits with a long line of research according to which humans encode complex concepts as nested combinations of a finite set of elementary primitives (Chomsky, 1956; Fodor, 1975; Hauser et al., 2002; Leeuwenberg, 1969; Restle, 1970). This idea was initially supported by behavioral studies of human rule learning (Shepard et al., 1961). In concept learning experiments, the speed and efficiency of learning was shown to be modulated by the Boolean complexity of the rule (i.e., the length of its shortest logical expression as a formula with elementary logical operators and parentheses) (Feldman, 2000). Working memory for sequences of digits was also found to be modulated by the presence of simpler chunks (Mathy and Feldman, 2012). Such research led to the general proposal that the human brain acts as a compressor of information in all sorts of domains and always attempts to select the shortest expression that accounts for what it perceives (Chater and Vitányi, 2003; Feldman, 2000, 2003; Li and Vitányi, 1993; Romano et al., 2013). The present research adds further evidence in favor of this framework. Indeed, the same formal language of geometry that we tested here was recently found to capture the regularities in binary auditory and visual sequences made of two arbitrary sounds or pictures (Planton et al., 2021). Future research should examine three key open questions. First, to what extent can the same set of recursive language-like rules capture very different domains in which humans excel, such as mathematics, music, and language? Second, how are such rules implemented at the neural level? And third, are such codes uniquely developed in human, as postulated by some researchers (Dehaene et al., 2015; Fitch, 2014; Hauser et al., 2002) or can they also be observed in non-human primates?

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Experimental protocol
  - MEG task
  - MEG acquisition and preprocessing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Time-resolved multivariate decoding
  - Decoding of spatial position
  - Anticipation score
  - Modulation of anticipation and syntactic role
  - Decoding primitive identity and ordinal position
  - Fourier analysis and presence of a peak
  - Representational similarity analysis
  - Statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2021.06.009.

ACKNOWLEDGMENTS

We acknowledge help from all the NeuroSpin support teams, particularly Leila Azizi and Virginie van Wassenhove. We are grateful to Maxime Maheu, Pedro Pinheiro-Chagas, and Darinka Trübutschek for their valuable comments. This research was supported by INSERM, CEA, Collège de France, the Bettencourt-Schueller Foundation, and a European Research Council (ERC) grant (Neurosyntax) to S.D. The work was performed on a platform of the France Life Imaging network, partly funded by Agence Nationale de la Recherche grant ANR-11-INBS-0006, and supported by the Bettencourt-Schueller Foundation and the Leducq Foundation. This article is dedicated to the memory of our friend and colleague Sébastien Marti, who directed our MEG center and died while this work was being completed.

AUTHOR CONTRIBUTIONS

Conceptualization, F.A.R., L.W., M.A., and S.D.; data curation, F.A.R. and S.M.; formal analysis, F.A.R.; validation, F.A.R. and S.D.; investigation, F.A.R. and S.M.; visualization, F.A.R.; methodology, F.A.R., L.W., and M.A.; writing – original draft, F.A.R.; writing – review & editing, S.D.; resources, S.M.; supervision, S.D.; funding acquisition, S.D.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 16, 2020
Revised: November 3, 2020
Accepted: June 7, 2021
Published: July 5, 2021

SUPPORTING CITATIONS

The following references appear in the supplemental information: Mostert et al., 2018, Quax et al., 2019.

REFERENCES

Amalric, M., and Dehaene, S. (2017). Cortical circuits for mathematical knowledge: evidence for a major subdivision within the brain's semantic networks. Philos. Trans. R. Soc. Lond. B Biol. Sci. *373*, 20160515.

Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., and Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. PLoS Comput. Biol. *13*, e1005273.

Anderson, J.R., Bothell, D., Lebiere, C., and Matessa, M. (1998). An integrated theory of list memory. J. Mem. Lang. *38*, 341–380.

Baddeley, A. (2003). Working memory: looking back and looking forward. Nat. Rev. Neurosci. *4*, 829–839.

Baddeley, A.D., and Hitch, G. (1974). Working memory. In Recent Advances in Learning and Motivation, *Vol. 8*, G.A. Bower, ed. (Academic Press), pp. 47–89.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn. Sci. *12*, 193–200.

Badre, D., and D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? Nat. Rev. Neurosci. *10*, 659–669.

Badre, D., and Nee, D.E. (2018). Frontal cortex and the hierarchical control of behavior. Trends Cogn. Sci. *22*, 170–188.

Badre, D., Kayser, A.S., and D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. Neuron *66*, 315–326.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical microcircuits for predictive coding. Neuron *76*, 695–711.

Bor, D., Duncan, J., Wiseman, R.J., and Owen, A.M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. Neuron *37*, 361–367.

Botvinick, M.M., and Plaut, D.C. (2006). Short-term memory for serial order: a recurrent neural network model. Psychol. Rev. *113*, 201–233.

Botvinick, M., and Watanabe, T. (2007). From numerosity to ordinal rank: a gain-field model of serial order representation in cortical working memory. J. Neurosci. *27*, 8636–8642.

Chao, Z.C., Takaura, K., Wang, L., Fujii, N., and Dehaene, S. (2018). Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. Neuron *100*, 1252–1266.e3.

Chater, N., and Vitányi, P. (2003). Simplicity: a unifying principle in cognitive science? Trends Cogn. Sci. *7*, 19–22.

Chomsky, N. (1956). Three models for the description of language. IEEE Trans. Inf. Theory *2*, 113–124.

Dehaene, S., Izard, V., Pica, P., and Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. Science *311*, 381–384.

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., and Pallier, C. (2015). The neural representation of sequences: from transition probabilities to algebraic patterns and linguistic trees. Neuron *88*, 2–19.

Demarchi, G., Sanchez, G., and Weisz, N. (2019). Automatic and feature-specific prediction-related neural activity in the human auditory system. Nat. Commun. *10*, 3440.

Desrochers, T.M., Chatham, C.H., and Badre, D. (2015). The necessity of rostrolateral prefrontal cortex for higher-level sequential behavior. Neuron *87*, 1357–1368.

Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. *19*, 158–164.

Ekman, M., Kok, P., and de Lange, F.P. (2017). Time-compressed preplay of anticipated events in human primary visual cortex. Nat. Commun. *8*, 15276.

Elman, J.L. (1990). Finding structure in time. Cogn. Sci. *14*, 179–211.

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. Nature *407*, 630–633.

Feldman, J. (2003). The simplicity principle in human concept learning. Curr. Dir. Psychol. Sci. *12*, 227–232.

Ferrigno, S., Cheyette, S.J., Piantadosi, S.T., and Cantlon, J.F. (2020). Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. Science Advances *6*, eaaz1002.

Fias, W., Lammertyn, J., Caessens, B., and Orban, G.A. (2007). Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulcus. J. Neurosci. *27*, 8952–8956.

Fitch, W.T. (2004). Computational Constraints on Syntactic Processing in a Nonhuman Primate. Science *303*, 377–380.

Fitch, W.T. (2014). Toward a computational framework for cognitive biology: unifying approaches from cognitive neuroscience and comparative cognition. Phys. Life Rev. *11*, 329–364.

Fodor, J.A. (1975). The Language of Thought (Harvard University Press).

Friston, K.J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage *19*, 1273–1302.

Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. Front. Neurosci. *7*, 267.

Hauser, M.D., Chomsky, N., and Fitch, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? Science *298*, 1569–1579.

Hurlstone, M.J., Hitch, G.J., and Baddeley, A.D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. Psychol. Bull. *140*, 339–373.

Jas, M., Larson, E., Engemann, D.A., Leppäkangas, J., Taulu, S., Hämäläinen, M., and Gramfort, A. (2018). A reproducible MEG/EEG group study with the MNE software: recommendations, quality assessments, and good practices. Front. Neurosci. *12*, 530.

Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., and Wang, L. (2018). Production of Supra-regular Spatial Sequences by Macaque Monkeys. Current Biology *28*, 1851–1859.e4.

King, J.R., and Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. Trends Cogn. Sci. *18*, 203–210.

Koechlin, E., and Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. Neuron *50*, 963–974.

Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science *302*, 1181–1185.

Kok, P., Jehee, J.F., and de Lange, F.P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. Neuron *75*, 265–270.

Kok, P., Failing, M.F., and de Lange, F.P. (2014). Prior expectations evoke stimulus templates in the primary visual cortex. J. Cogn. Neurosci. *26*, 1546–1554.

Kok, P., Mostert, P., and de Lange, F.P. (2017). Prior expectations induce prestimulus sensory templates. Proc. Natl. Acad. Sci. U S A *114*, 10473–10478.

Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.

Kutter, E.F., Bostroem, J., Elger, C.E., Mormann, F., and Nieder, A. (2018). Single neurons in the human brain encode numbers. Neuron *100*, 753–761.e4.

Leeuwenberg, E.L. (1969). Quantitative specification of information in sequential patterns. Psychol. Rev. *76*, 216–220.

Lewandowsky, S., and Murdock, B.B., Jr. (1989). Memory for serial order. Psychol. Rev. *96*, 25–57.

Li, M., and Vitányi, P. (1993). An Introduction to Kolmogorov Complexity and Its Applications (Springer).

Liu, Y., Dolan, R.J., Kurth-Nelson, Z., and Behrens, T.E.J. (2019). Human replay spontaneously reorganizes experience. Cell *178*, 640–652.e14.

Marcus, G.F., Vijayan, S., Bandi Rao, S., and Vishton, P.M. (1999). Rule learning by seven-month-old infants. Science *283*, 77–80.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods *164*, 177–190.

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

Mostert, P., Albers, A.M., Brinkman, L., Todorova, L., Kok, P., and de Lange, F.P. (2018). Eye movement-related confounds in neural decoding of visual working memory representations. eNeuro *5*, ENEURO.0401-17.2018.

Nelson, M.J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., Cash, S.S., Naccache, L., Hale, J.T., Pallier, C., and Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. Proc. Natl. Acad. Sci. U S A *114*, E3669–E3678.

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. Proc. Natl. Acad. Sci. U S A *109*, 11860–11865.

Nieder, A., and Dehaene, S. (2009). Representation of number in the brain. Annu. Rev. Neurosci. *32*, 185–208.

Nieder, A., Diester, I., and Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. Science *313*, 1431–1435.

Ninokura, Y., Mushiake, H., and Tanji, J. (2004). Integration of temporal order and object information in the monkey lateral prefrontal cortex. J. Neurophysiol. *91*, 555–560.

Orlov, T., Yakovlev, V., Hochstein, S., and Zohary, E. (2000). Macaque monkeys categorize images by their ordinal number. Nature *404*, 77–80.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

Planton, S., van Kerkoerle, T., Abbih, L., Maheu, M., Meyniel, F., Sigman, M., Wang, L., Figueira, S., Romano, S., and Dehaene, S. (2021). A theory of memory for binary sequences: evidence for a mental compression algorithm in humans. PLoS Comput. Biol. *17*, e1008598.

Quax, S.C., Dijkstra, N., van Staveren, M.J., Bosch, S.E., and van Gerven, M.A.J. (2019). Eye movements explain decodability during perception and cued attention in MEG. Neuroimage *195*, 444–453.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. Psychol. Rev. *77*, 481–495.

Romano, S., Sigman, M., and Figueira, S. (2013). LT2C2: a language of thought with Turing-computable Kolmogorov complexity. Pap. Phys. *5*, 050001.

Romano, S., Salles, A., Amalric, M., Dehaene, S., Sigman, M., and Figueira, S. (2018). Bayesian validation of grammar productions for the language of thought. PLoS ONE *13*, e0200420.

Sakai, K., and Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. Nature *354*, 152–155.

Shepard, R.N., Hovland, C.I., and Jenkins, H.M. (1961). Learning and memorization of classifications. Psychol. Monogr. *75*, 1.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. Artif. Intell. *46*, 159–216.

Smolensky, P., and Legendre, G. (2006). The Harmonic Mind (MIT Press).

Spelke, E., Lee, S.A., and Izard, V. (2010). Beyond core knowledge: natural geometry. Cogn. Sci. *34*, 863–884.

Summerfield, C., and de Lange, F.P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. Nat. Rev. Neurosci. *15*, 745–756.

Taulu, S., Kajola, M., and Simola, J. (2004). Suppression of interference and artifacts by the signal space separation method. Brain Topogr. *16*, 269–275.

Terrace, H.S., Son, L.K., and Brannon, E.M. (2003). Serial expertise of rhesus macaques. Psychol. Sci. *14*, 66–73.

Wang, L., Amalric, M., Fang, W., Jiang, X., Pallier, C., Figueira, S., Sigman, M., and Dehaene, S. (2019). Representation of spatial sequences using nested rules in human prefrontal cortex. Neuroimage *186*, 245–255.

Wang, L., Uhrig, L., Jarraya, B., and Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. Current Biology *25*, 1966–1974.

Yildirim, I., and Jacobs, R.A. (2015). Learning multisensory representations for auditory-visual transfer of sequence category knowledge: a probabilistic language of thought approach. Psychon. Bull. Rev. *22*, 673–686.

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
MEG data	This paper	N/A
Software and algorithms		
MATLAB	MathWorks	https://www.mathworks.com/products/matlab.html
PYTHON	Anaconda	https://www.anaconda.com/
Custom code and algorithms	This paper	https://github.com/Fosca/GeomSeq
Other		
Neural recordings and amplifier	306-channel, whole-head MEG by Elekta Neuromag®	https://www.elekta.com
Eye Tracking Device	Eyelink 1000	https://www.sr-research.com/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Fosca Al Roumi (fosca.alroumi@gmail.com).

Materials availability

The study did not produce new materials.

Data and code availability

The original data supporting the current study has not been deposited in a public repository. The code is available at https://github.com/Fosca/GeomSeq.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

20 participants (9 men, M_{age} = 24.6 years, SD_{age} = 3.7 years) with normal vision were included in the MEG experiment. We didn't test any effect of gender on the results of this study. In compliance with institutional guidelines, all subjects gave written informed consent prior to enrollment and received 90€ as compensation.

METHOD DETAILS

Experimental protocol

General structure of the experiment

The main task, completed in the MEG Elekta acquisition device, was subdivided into 3 parts. To avoid biasing subjects toward specific primitives, the sequence part was performed first. The second part was dedicated to each of the primitive operations, and the third part was a localizer task with unpredictable locations, designed to train a decoder for spatial locations. During the 3 parts of the MEG experiment, white dots were flashed for 100ms on the vertices of an octagon while the subject was fixating a cross at the center of the screen. The MEG experiment was preceded by a short training (c.a. 20 minutes) to the geometrical sequences.

Training on sequences

As initial training, outside the MEG, the geometrical sequences were presented with a slower pace than the rest of the experiment: a stimulus onset asynchrony (SOA) of 700ms between consecutive dots, and a dot duration of 200ms. Each sequence was repeated until the participants pressed the space bar to report that they had memorized it. They were then asked to type in the 8 locations that followed the last item displayed on the screen. If they succeeded, the word 'Bravo' (congratulations) was presented on the screen and the next sequence started. If they failed, the word 'Erreur' (error) was displayed and the same sequence restarted. Participants were instructed that the same sequences would be presented to them during the main experiment. No training was provided on the primitive part of the experiment. The training part was meant to select participants able to quickly encode the geometrical sequences. Only the ones that had finished the training in less than 20 minutes were qualified for the main MEG experiment. 5 out of 25 participants did not manage to do the training part of the experiment in less than 20 minutes.

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

MEG task

Geometrical sequences

The first part of the experiment was devoted to the geometrical sequences and was composed of 4 runs. 9 sequences (Figure 1B) were composed of 8 non-repeating locations. The last 3, called 'Memory sequences', which were composed of 1, 2 or 4 spatial locations, were not analyzed in this study. During one run, each sequence of 8 locations was presented 12 times consecutively. The sequences presented in Figure 1B were mere templates: to generate the actual sequences, the starting point and global direction of rotation were varied and balanced across runs. 4segment sequences were selected such that each of the 4 symmetry axes (horizontal, vertical and diagonal) would appear once. Participants had to perform two tasks. First, they had to report with a button press when they had identified the sequence and felt able to predict the next locations. Second, during the 11^{th} or the 12^{th} repetition, an item appeared at an unexpected location, and subjects had to report this violation with another button press as fast as possible. This task was added to maintain participants attention during the full block. MEG epochs containing such a violation were excluded from all analyses.

Primitive part

The second part of the experiment was devoted to the primitive operations. It was composed of 4 runs, subdivided in 12 mini-blocks for each of the 12 conditions. The 11 first ones followed elementary 'primitive' rules (Figure 1C), because a simple geometrical operation allowed to determine the spatial location of the second item of a pair when given the first. The 12^{th} condition was a control condition in which no minimal rule allowed to do so, and participants could only memorize the 8 unrelated pairs in order to perform the task. A mini-block was composed of 32 pairs with a SOA of 433ms between the items of the pair and an inter-pair-interval of 1100ms. Each of the 8 pairs appearing 4 times in the mini-block.

The task was similar to the sequence part. Participants reported with a button press when they had identified the rule that allowed them to predict the location of the second item given the first. In addition, they had to detect as fast as possible when the second item did not appear at the expected location. Violations could only occur during the presentation of the last 8 items of the mini-block. Again, MEG epochs containing such a violation were excluded from all analyses.

Localizer part

The last part of the experiment was meant to train a decoder for spatial position of the presented items. To do so, dots were flashed pseudo-randomly on the vertices of the octagon with SOA 433ms. Occasionally (1/20 dots on average) the color of the dot changed. The subject had to click as fast as possible to report this.

MEG acquisition and preprocessing

MEG recordings

Participants performed the tasks while sitting inside an electromagnetically shielded room. The magnetic component of their brain activity was recorded with a 306-channel, whole-head MEG by Elekta Neuromag® (Helsinki, Finland). 102 triplets, each comprising one magnetometer and two orthogonal planar gradiometers composed the MEG helmet. The brain signals were acquired at a sampling rate of 1000 Hz with a hardware highpass filter at 0.03Hz. The data was then decimated by a factor 4.

Eye movements and heartbeats were monitored with vertical and horizontal electro-oculograms (EOGs) and electrocardiograms (ECGs). Subjects' head position inside the helmet was measured at the beginning of each run with an isotrack Polhemus Inc. system from the location of four coils placed over frontal and mastoïdian skull areas.

Data cleaning: Maxfiltering

Bad MEG channels were identified visually in the raw signal and were provided to the MaxFilter software (ElektaNeuromag®, Helsinki, Finland) to compensate for head movements between experimental blocks by realigning all data to an average head position and to apply the signal space separation algorithm (Taulu et al., 2004) to suppress magnetic interference from outside the sensor helmet and interpolate bad channels.

Data cleaning: ICA

The rest of the analysis was performed with MNE Python (Gramfort et al., 2013; Jas et al., 2018). Oculomotor and cardiac artifacts were removed performing an independent component analysis (ICA). The components that correlated the most with the EOG and ECG signals were automatically detected. We then visually inspected their topography and correlation to the ECG and EOG time series to confirm their rejection from the MEG data.

Recording participants' gaze

During the MEG acquisition, participants were instructed to fixate the central cross while the items were flashed. Their gaze was monitored online to make sure that they did. Eye-tracking data was collected for 14 out of 20 participants using EyeLink 1000 eye-tracker device (SR research).

QUANTIFICATION AND STATISTICAL ANALYSIS

Time-resolved multivariate decoding

The goal multivariate of time-resolved decoding analyses was to predict from single-trial brain activity (X) a specific categorical (e.g., primitive identity) or continuous (e.g., angular position) variable (y) that represented the neuronal state corresponding to the participant's mental representation. These analyses were performed following King et al.'s preprocessing pipeline (King and Dehaene, 2014) using

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

MNE-python (Gramfort et al., 2013). Prior to model fitting, each channel at each time-point was z-scored across trials. Each estimator was fitted on each participant separately, across all MEG sensors using the parameters set to their default values provided by the Scikit-Learn package (Pedregosa et al., 2011). When the estimator was trained and tested on two different conditions, the whole training and testing sets were used to respectively fit and test the estimator. By contrast, when the decoder was trained and tested on non-independent data, we used a stratified cross-validation procedure with 5 folds for spatial decoding, or cross-validated across block number for primitive and ordinal decoding. The reported scores are the average across cross-validation folds.

Decoding of spatial position

Data from a localizer block, where locations were randomly intermixed, as well as data from the first item of each pair in the pair block, which was unpredictable was used to train spatial angular decoder.

The spatial angular decoders were built from two ridge regressions used to decode the angular position Θ. One predicted $\sin(\Theta)$ and the other $\cos(\Theta)$. The angular decoding score was obtained by first computing the mean absolute difference between the predicted angle (Θ_{pred}) and the true angle (Θ_{true}). We subtracted to this score $\pi/2$ to obtain a score in the range of $-\pi/2$ and $\pi/2$ (chance = 0) (King and Dehaene, 2014).

To access the temporal organization of the neural representations, we computed the generalization-across-time (GAT) matrices. These matrices represent the decoding score of an estimator trained at time t (training time on the vertical axis) and tested with data from another time t' (testing time on the horizontal axis).

Anticipation score

The anticipation score was obtained from the predictions of an angular decoder trained on brain data decimated by a factor of 2, averaged over windows of 10ms every 5ms. This was meant to increase the signal-to-noise ratio while avoiding the contamination of anticipation signals by the visual response to the expected item. The decoder was then tested on the sequence data. However, we removed the first presentation of the sequence, that was needed to identify it, and the two last repetitions, as a violation could occur. The output of the decoder was binned into the 8 spatial positions. The anticipation score was defined as the subtraction of the proportion of times the correct position was predicted and the proportion of times the other position at the same distance from the preceding item was predicted (Figure 4). This measure was designed to overcome a potential confound that comes from the fact that successive sequence items tend to be close to each other (average angle between two sequence items is 73°, s.d. = 28°). As the angular decoder output spreads on neighboring angles, it may predict with an above chance performance the spatial position of the next item. This is particularly true for simpler sequences, which involve shorter distances. Therefore, by construction, the anticipation score is immune to such a distance-based confound.

The anticipation score is not defined when the next item is at distance 4 from the previous one. As this represents 50% of the transitions for 4diagonals and 2crosses sequences, we excluded these sequences from the analyses. We ran linear regressions per participant on the average of the anticipation scores across sequence types. The statistics are computed on the subjects' distribution of regression coefficients.

Modulation of anticipation and syntactic role

We compared the anticipation scores of the items that, according to our postulated language, open a component, to the ones that are inside a component. The analysis was run only on the 4segments and 2squares sequences. The repeat, alternate and irregular sequences were excluded from it as their postulated representation does not involve a nested syntactic structure. We also excluded the 2arcs sequences since all constituent opening corresponded to distance-4 transitions. Finally, we also discarded the complex 2rectangle sequence, since its anticipation score was not significantly different from zero.

Decoding primitive identity and ordinal position

The data was smoothed over sliding windows of 100ms for every time-step (i.e., each 4ms). We used Support Vector Machines (SVM) to decode which geometrical primitive was involved at a given transition between two locations, as predicted by our language model. When controlling for visual confounds, rotation ± 2 primitive was excluded from the trial set as no symmetries involved a distance of 2. Moreover, when only two categories were considered (e.g., 'rotation' and 'symmetry'), the score of the decoder was provided in terms of the area under the curve. When decoding the 11 primitives, we used One-VS-rest multiclass classifiers and report their mean accuracy.

A four-way decoder (One-VS-rest multiclass classifier) was used to decode ordinal positions 1-4; and a binary decoder was used to decode ordinal positions 1 versus 2. This ordinal position was at the lowest level of the hierarchical code predicted by our language model (i.e., 1-2-1-2… or 1-2-3-4 depending on the sequence, see Figure 6). The decoders' performance was estimated by computing their mean accuracy, averaged over participants.

Significance was assessed by a cluster-based permutation test on the window −200 to 600ms.

Fourier analysis and presence of a peak

To determine the periodicity of the activation of the ordinal code, we ran a fast Fourier transform on the projection on the decision axes time-series using *scipy*'s *fftpack* module. We computed the difference between the log-power at the test frequency and the

Please cite this article in press as: Al Roumi et al., Mental compression of spatial sequences in human working memory using numerical and geometrical primitives, Neuron (2021), https://doi.org/10.1016/j.neuron.2021.06.009

log-power at neighboring frequencies. The log-power at f (2.31 Hz), f/2 (1.15 Hz) and f/4 (0.577 Hz) was compared to the average log-power for the 4 neighboring frequencies (frequency bin width: 0.024 Hz) using a Student's t test. To determine if there was an interaction between the amplitude at f/2 and f/4 and the component size, we ran a two-factor within subjects repeated-measures ANOVA with MATLAB®.

Representational similarity analysis

Representational similarity analysis (RSA) characterizes a neural representation by the similarity between the neural patterns elicited by a set of stimuli. To test a given model, we compare the representational similarity it predicts to the one measured from brain signals. In this experiment, we hypothesized that brain activity would reflect a superposition of visuospatial factors (item location, distance between consecutive items) and high-level geometrical primitives. To compute the similarity between two conditions, we smoothed the data over sliding windows of 100ms with 10ms steps. Epochs belonging to the even and odd run numbers were averaged separately to form two sets of evoked activities. The empirical similarity was determined as the Spearman rank correlation between the evoked activities of these two sets. The empirical dissimilarity (1-Spearman rank correlation) was then regressed as a function of the theoretical representational dissimilarities. These predictor matrices were z-scored beforehand.

Statistical analyses

All statistics reported in the text refer to group-level analyses. Tukey post hoc tests were performed using the R software packages *nlme* and *multcomp*. We used permutation statistics to assess multivariate decoding performance for the two-dimensional time by time generalization-across-time (GAT) matrices and for the simple decoding score time-courses. We considered temporal clusters and non-parametric one-sample t tests estimated on 4096 permutations (Maris and Oostenveld, 2007) implemented by the *permutation_cluster_1samp_*test available in *mne.stats* package. A cluster was defined by adjacent time points. The cluster-level statistic was the sum of the sample-specific *t*-statistics that belonged to a given cluster. The alpha level of the sample-specific test statistic and of the cluster-specific test statistic were 0.05. For decoding performance curves, '*' and '**' indicate that resulting p values are respectively < 0.05 and < 0.01. Dashed contours on temporal generalization correspond to p value < 0.05 resulting from the permutation test. The Student's t tests and the linear regressions were performed with MATLAB®. Stepwise linear regressions were computed with the MATLAB® *stepwiselm* function using an AIC criterion starting from a model with intercept. *rm_anova2* function was used to run the two-factor within subjects repeated-measures ANOVA with MATLAB®, available on the MATLAB® File Exchange.