# Compression of binary sound sequences in human working memory

Samuel Planton[1*], Fosca Al Roumi[1*+], Liping Wang[2] & Stanislas Dehaene[1,3]

[1]Cognitive Neuroimaging Unit, Université Paris-Saclay, INSERM, CEA, CNRS, NeuroSpin center, 91191 Gif/Yvette, France
[2]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
[3]Collège de France, Université Paris Sciences Lettres (PSL), 11 Place Marcelin Berthelot, 75005 Paris, France

*Co-first authors
+ Corresponding author
Postal address:
Cognitive Neuroimaging Unit, INSERM-CEA-University Paris Saclay NeuroSpin center, CEA/SAC/DRF/Joliot Bât 145, Point Courrier 156 F-91191 Gif/Yvette, France
E-mail: fosca.alroumi@gmail.com

# Abstract

According to the language of thought hypothesis, regular sequences are compressed in human working memory using recursive loops akin to a mental program that predicts future items. We tested this theory by probing working memory for 16-item sequences made of two sounds. We recorded brain activity with functional MRI and magneto-encephalography (MEG) while participants listened to a hierarchy of sequences of variable complexity, whose minimal description required transition probabilities, chunking, or nested structures. Occasional deviant sounds probed the participants' knowledge of the sequence. We predicted that task difficulty and brain activity would be proportional to minimal description length (MDL) in our formal language. Furthermore, activity should increase with MDL for learned sequences, and decrease with MDL for deviants. These predictions were upheld in both fMRI and MEG, indicating that sequence predictions are highly dependent on sequence structure and become weaker and delayed as complexity increases. The proposed language recruited bilateral superior temporal, precentral, anterior intraparietal and cerebellar cortices. These regions overlapped extensively with a localizer for mathematical calculation, and much less with spoken or written language processing. We propose that these areas collectively encode regular sequences as repetitions with variations and their recursive composition into nested structures.

# Introduction

The ability to learn and manipulate serially ordered lists of elements, i.e. sequence processing, is central to several human activities (Lashley, 1951). This capacity is inherent to the ordered series of subtasks that make up the actions of daily life, but is especially decisive for the implementation of high-level human skills such as language, mathematics, or music. In non-human primates, multiple levels of sequence encoding ability, with increasing complexity, have been identified, from the mere representation of transition probabilities and timings to ordinal knowledge (which element comes first, second, third…), recurring chunks, and even abstract patterns (e.g. does the sequence obey the pattern xxxxY, i.e a repetition ending with a different element) (Dehaene et al., 2015; Jiang et al., 2018; Shima et al., 2007; Wang et al., 2015; Wilson et al., 2013). We and others, however, proposed that the representation of sequences in humans may be unique in its ability to encode recursively nested hierarchical structures, similar to the nested phrase structures that linguists postulate to underlie human language (Dehaene et al., 2015; Fitch & Martins, 2014; Hauser et al., 2002). Building on this idea, it was suggested that humans would spontaneously encode temporal sequences of stimuli using a language-like system of nested rules, a "language of thought" (LoT) (Fodor, 1975) (Al Roumi et al., 2021; Amalric et al., 2017; Chater & Vitányi, 2003; Feldman, 2000; Li & Vitányi, 1993; Mathy & Feldman, 2012; Planton et al., 2021; Wang et al., 2019). For instance, when faced with a sequence such as xxYYxYxY, humans may encode it using an abstract internal expression equivalent to « 2 groups of 2, and then an alternation of 4 ».

The assumption that humans encode sequences in a recursive, language-like manner, was recently tested with a non-linguistic visuo-spatial task, by asking human adults and children to memorize and track geometric sequences of locations on the vertices of an octagon (Al Roumi et al., 2021; Amalric et al., 2017; Wang et al., 2019). Behavioral and brain-imaging studies showed that such sequences are internally compressed in human memory using an abstract "language of geometry" that captures their numerical and geometrical regularities (e.g., "next element clockwise", "vertical symmetry"). Indeed, behavioral results showed that the difficulty of memorizing a sequence was linearly modulated, not by the actual sequence length, but by the length of the program capable of generating it using the proposed language ("minimal description length" or MDL; for a definition and brief review, see Dehaene et al., 2022). In a follow-up fMRI experiment where participants had to follow the same sequences with their gaze, activity in the dorsal part of inferior prefrontal cortex correlated with the LoT-complexity while the right dorsolateral prefrontal cortex (dlPFC) encoded the presence of embedded structures. These results indicate that sequences are stored in memory in a compressed manner, the size of this code being the length of the shortest program that describes the sequence in the proposed formal language. Working memory for sequences would therefore follow the “minimal description length” principle inherited from information theory (Grunwald, 2004) and often used to capture various human behavior (Chater & Vitányi, 2003; Feldman, 2000; Mathy & Feldman, 2012). Wang et al. (2019) further showed that the encoding and compression of such sequences involved brain areas supporting the processing of mathematical expressions rather than language-related areas, suggesting that multiple internal languages, not necessarily involving classical language areas, are present in the human brain. In a follow-up study, Al Roumi et al. (2021) showed with MEG that the spatial, ordinal, and geometrical primitive codes postulated in the proposed LoT could be extracted from brain activity.

In the present work, we ask whether this LoT may also explain the human memory for binary auditory sequences (i.e. sequences made up of only two possible items, for instance two sounds with high and low pitch, respectively). While arguably minimal, binary sequences preserve the possibility of forming structures at different hierarchical levels. They therefore provide an elementary window into the mental representation of nested language-like rules, and which aspect of this representation, if any, is unique to the human species. While it would make little sense to ask if non-human animals can store spoken human sentences, it does seem more reasonable to submit them to a protocol with minimal, binary sound sequences, and ask whether they use a recursive language-like format for encoding in working memory, or whether they are confined to statistical learning or chunking. The latter mechanisms are important to consider because they are thought to underpin the processing of several aspects of sequence processing in human infants and adults as well as several animal species, such as the extraction of chunks within a stream of syllables, tones or shapes (Wang et al., 2019), or the community structure that generates a sequence of events (Karuza et al., 2019; Schapiro et al., 2013). Yet, very few studies have tried to separate the brain mechanisms underlying rule-based predictions from those of probabilistic sequence learning (Bhanji et al., 2010; Kóbor et al., 2018; Maheu et al., 2021). Our goal here is to develop such a paradigm in humans, and to test the hypothesis that human internal models are based on a recursive language of thought.

The present work capitalized on a series of behavioral experiments (Planton et al., 2021), we recently proved that human performance in memorizing binary auditory sequences, as tested by the capacity to detect occasional violations, could be predicted by a modified version of the language of geometry, based on the hierarchical combination of very few primitives (repeat, alternate, concatenate, and integers). This work considered binary sequences of various lengths (from 6 to 16 items) mainly in the auditory but also in the visual modality, and showed that MDL in such a formal language accurately predicted participants' oddball detection performance. This was especially true for longer sequences of 16 items as their length exceed typical working memory capacity (Cowan, 2001, 2010; Miller, 1956). In this work, LoT predictions were compared to competitor models of cognitive complexity and information compression (Aksentijevic & Gibson, 2012; Alexander & Carey, 1968; Delahaye & Zenil, 2012; Gauvrit et al., 2014; Glanzer & Clark, 1963; Psotka, 1975; Vitz & Todd, 1969). The predictive power of LoT outperformed all competing theories (Planton et al., 2021).

Here, we use functional MRI and magneto-encephalography to investigate the cerebral underpinnings of the proposed language in the human brain. We exposed participants to 16-item auditory binary sequences, with varying levels of regularity, while recording their brain activity with fMRI and MEG in two separate experiments (see Figure 1). By combining these two techniques, we aimed at obtaining both the spatial and the temporal resolution needed to characterize in depth the neural mechanisms supporting sequence encoding and compression.

In both fMRI and MEG, the experiment was composed of two phases. In a habituation phase, the sequences were repeatedly presented in order for participants to memorize them, thus probing the complexity of their internal model. In a test phase, sequences were occasionally presented with deviants (a single tone A replacing another tone B), thus probing the violations of expectations generated by the internal model (Figure 1B). We focused on a very simple prediction arising from the hierarchical predictive coding framework (Friston, 2005). According to this view, and to much experimental research (Bekinschtein et al., 2009; e.g. Chao et al., 2018; Heilbron & Chait, 2018; Summerfield & de Lange, 2014; Wacongne et al., 2011), the internal model of the sequence, as described by the postulated LoT, would be encoded by prefrontal regions of the brain, and would send anticipation signals to auditory areas, where they would be subtracted from incoming signals. As a consequence, we predict a reciprocal effect of LoT on the brain signals during habituation and during deviancy. In the habituation part of the experiment, lower amplitude response signals should be observed for sequences of low complexity – and conversely, during low complexity sequences, we expect top-down predictions to be stronger and therefore deviants to elicit larger responses, than for complex, hard to predict sequences.

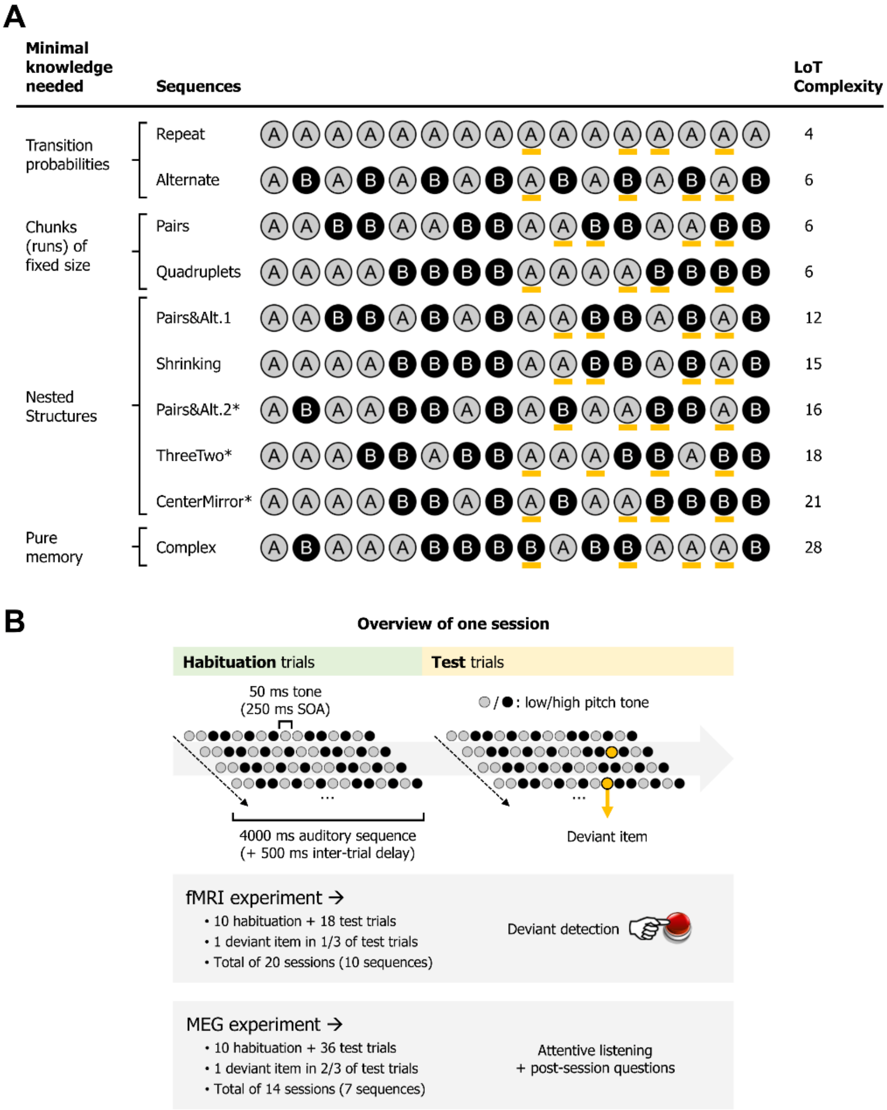


**Figure 1. Experimental design probing sequence knowledge in humans**. **A)** List of sequences used in MEG and fMRI experiments. All sequences comprised 16 occurrences of the same two sounds (low or high pitch, here depicted as A and B). Sequences formed a hierarchy of complexity in the proposed language of thought (LoT, right column). They were categorized according to the minimal type of knowledge assumed to be required for optimal memory encoding (left column). Orange lines indicate the positions at which violations could occur. Stars (*) show sequences used only in the fMRI experiment. **B)** Overview of the presentation paradigm: In each session, participants were first presented with several repetitions of a fixed sequence (habituation period), then their working memory for that sequence was tested with occasional deviant probes. Each sequence was tested twice in two different sessions, while reversing the mapping between A/B items and low/high pitch.

# Results

## Stimulus design

We designed a hierarchy of sequences (figure 1) of fixed length (16 items) that should systematically vary in complexity according to our previously proposed language of thought (Planton et al., 2021) and whose gradations separate the lower-level representations of sequences that may be accessible to non-human primates (as outlined in Dehaene et al., 2015a) from the more abstract ones that may only be accessible to humans (Figure 1A).

First, much evidence indicates that the brain spontaneously encodes statistical regularities such as transition probabilities in sequential sensory inputs and uses them to make predictions (e.g. Barascud et al., 2016; Bendixen et al., 2009; McDermott et al., 2013; Meyniel et al., 2016; Saffran et al., 1996), an ability well within the grasp of various non-human animals (e.g., Hauser et al., 2001; Meyer & Olson, 2011). The first two sequences in our hierarchy therefore consisted in predictable repetitions (AAAA…) and alternations (ABABA…). In terms of information compression, such sequences can be represented with a very short expression in our LoT model, in which repetitions or alternations are primitive operations out of which more complex expressions are built.

At the next level, we tested chunking, the ability to group a recurring set of contiguous items into a single unit, another major sequence encoding mechanism which is also accessible to non-human primates (Buiatti et al., 2009; Fujii & Graybiel, 2003; Saffran et al., 1996; Sakai et al., 2003; Uhrig et al., 2014). Thus, we included sequences made of pairs (AABBAABB…) or quadruplets (AAAABBBB…). Our LoT model attributes them a high level of compressibility, but already some degree of hierarchy (a loop of chunks). Relative to the previous sequenes, they require monitoring the number of repetitions before a new chunk starts (ABABA… = 1; AABBAA… = 2; AAAABBBB… = 4), and may therefore be expected to engage the number system, though to involve the bilateral intraparietal sulci, particularly their horizontal and anterior segments (Dehaene et al., 2003; Eger et al., 2009; Harvey et al., 2013; Kanayet et al., 2018).

The next level probed a more abstract level of sequence encoding, requiring nested structures, or a hierarchical representation of smaller chunks embedded in larger chunks. Although there is some debate on whether this level of representation could be accessed by non-human animals, especially with extensive training (Ferrigno et al., 2020; Gentner et al., 2006; Jiang et al., 2018; van Heijningen et al., 2009), many agree that the ability to access it quickly and spontaneously is a potential human-specific trait in sequence learning and its many related cognitive domains (Dehaene et al., 2015; Fitch, 2004; Fitch & Martins, 2014; Hauser et al., 2002). We probe it using a variety of complex but compressible sequences such as "AABBABABAABBABAB" (whose hierarchical description is $[A^2B^2[AB]^2]^2$ and can be paraphrased as "a repetition of two pairs followed by four alternations ). Here again, our LoT model easily compresses such nested structures by using only one additional bit whenever a chunk needs to be repeated, regardless of its hierarchical depth (for details, see Amalric et al., 2017).

Finally, as a control, our paradigm also includes a minimally compressible sequence, with balanced transition probabilities and minimal chunking possibilities. We selected a sequence which our language predicted to be of maximal complexity (highest MDL), and which was therefore predicted to challenge the limits of working memory (Figure 1A). Note that because such a complex sequences, devoid of recurring regularities, is not easily encodable within our language (except for a trivial concatenation of chunks), we may expect the brain areas involved in nested sequence coding to exhibit no further increase in activation, or even a decrease (Vogel & Machizawa, 2004). The presence of such a non-linear trend at the highest level of complexity may be tested by a quadratic contrast for MDL instead of a purely linear regression model.

## Behavioral data

After brain imaging, we asked all participants to report their intuitions of how each sequence could be parsed by drawing brackets on a visual representation of its contents (after listening to it). The results (see heatmaps in Figure 2A) indicated that participants agreed about how a sequence should be parsed and used bracketing levels appropriately for nested sequences. For instance, they consistently placed brackets in the middle of sequences that consisted in two phrases of 8 items, but did so less frequently both within those phrases and when the midpoint was not a predicted parsing point (sequences Pairs&Alt2 and CenterMirror in figure 2A). In order to assess the correspondence between the parsings and the organization proposed by the LoT model, we computed for each sequence the correlation between the group-averaged number-of-brackets vector and the LoT model vector (obtained from the sequence segmentation derived from the LoT description in terms of repeat, alternate and concatenate instructions). A strong correlation was found for sequences *Repeat* (Pearson r = 0.96, p < .0001), *Pairs* (r = 0.88, p < .0001), *Quadruplets* (r = 0.96, p < .0001), *Pairs&Alt.1* (r = 0.94, p < .0001), *Shrinking* (r = 0.93, p < .0001), *Pairs&Alt.2* (r = 0.85, p < .0001), *ThreeTwo* (r = 0.95, p < .0001), *CenterMirror* (r = 0.95, p < .0001) and *Complex* (r = 0.84, p < .0001), but not for *Alternate* (r = 0.08, p = .77). For the latter, a minor departure from the proposed encoding was found, as the shortest LoT representation (i.e. [+0]^16<b>) can be paraphrased as "16 alternations", while the participants' parses corresponded to "8 AB pairs". The latter encoding, however, only has a marginally larger complexity, so this deviation should not affect subsequent results.

Performance in the fMRI deviant detection task provided a more quantitative test of the model (similar to Planton et al., 2021). Sensitivity (d') was calculated by examining the hit rate for each sequence and each violation position, relative to the overall false-alarm rate on standard no-violation trials. On average, participants managed to detect the deviants at above chance level in all sequences and at all positions (Figure Sxxx; min d'= 0.556, min T(22) = 2.919, p < .0080). Thus, they were able to detect a great variety of violation types in regular sequences (unexpected alternations, repetitions, change in number, or chunk boundaries). However, performance worsened as the 16-item sequence became too complex to be easily memorized. The group-averaged performance in violation detection for each sequence (regardless of deviant position) was linearly predicted by LoT complexity, both for response times (RTs) (F(1, 8) = 43.87, p < .0002, $R^2$ = .85) and for sensitivity (d') (F(1, 8) = 159.4, p < .0001, $R^2$ = .95) (see Figure 2B). When including the participant as a random factor in a linear mixed model, we obtained a very similar result for sensitivity (F(1, 206) = 192.92, p < .0001, with estimates of -0.092 ± 0.007 for the LoT complexity predictor, and 3.39 ± 0.17 for the intercept), as well as for responses times (F(1, 203) = 110.87, p < .0001, with estimates of 17.4 ms ± 1.6 for the LoT complexity predictor, and 475.4 ms ± 38.6 for the intercept). As for false alarms, they were rare and no significant linear relationship was found in group averages (F(1, 8) = 2.18, p = .18), although a small effect was found in a linear mixed model with participant as the random factor (F(1, 206) = 4.83, p < .03, with estimates of 0.038 ± 0.017 for the LoT complexity predictor, and 1.57 ± 0.39 for the intercept).

We evaluated whether these results could be explained by statistical learning, i.e. whether deviants were more easily or more rapidly detected when they violated the transition probabilities of the current sequence. For sensitivity, a likelihood ratio test showed that adding a transition-probability measure of surprise (Maheu et al., 2019; Meyniel et al., 2016) to the linear regression with LoT complexity slightly improved the goodness of fit ($\chi^2(1) = 4.33$, $p < .038$). The effect of surprise was indeed significant in the new model ($F(1, 205) = 4.33$, $p < 0.039$), but the LoT complexity effect remained highly significant ($F(1, 205) = 106.40$, $p < .0001$). Similarly, for RTs, adding surprise to the model significantly improved model fit ($\chi^2(1) = 12.28$, $p < .0005$). Surprise explained some of the variance in RTs ($F(1, 202) = 12.53$, $p < .0005$), but the effect of LoT complexity remained highly significant ($F(1, 202) = 46.3$, $p < .0001$).

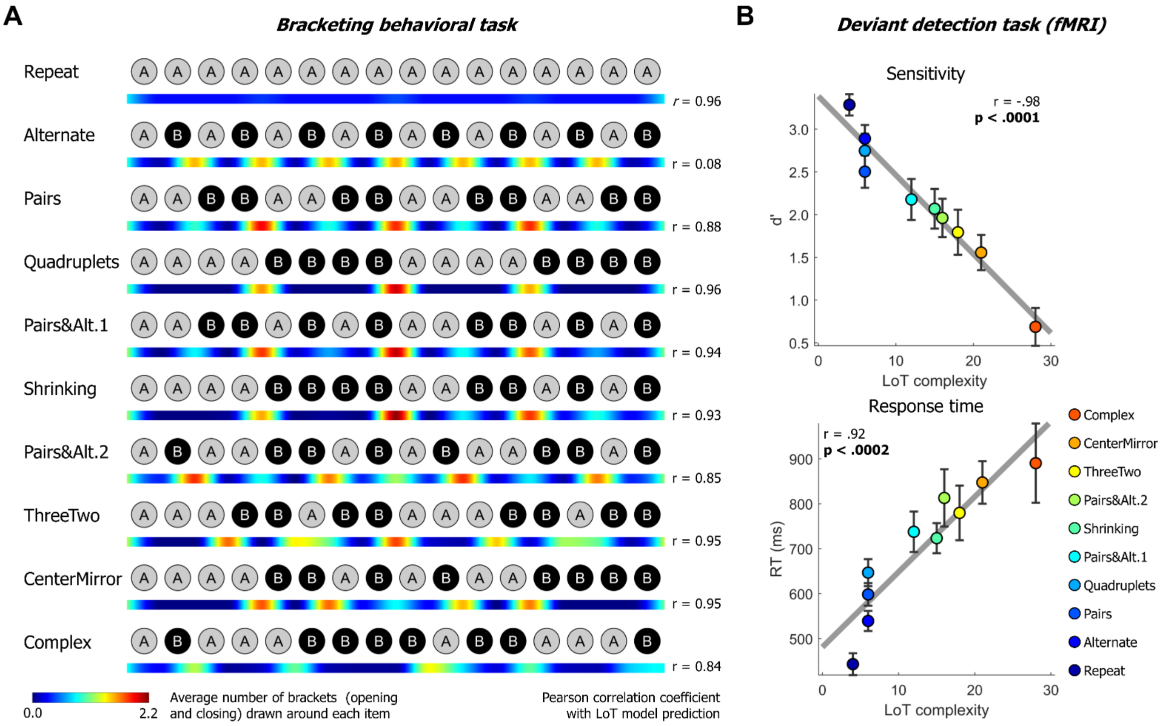


**Figure 2. Behavioral data supporting the existence of a recursive language of thought in humans**. **A)** Bracketing task. After the experiment, participants were asked to place brackets around a visual depiction of the sequence to depict how they mentally structured it. The heatmap for each sequence represent the average number of opening or closing brackets draw by the participants around each item (with smoothing for illustration purposes only). The Pearson correlation coefficient with the vector of brackets predicted by the LoT model is reported on the right side. A high correlation was obtained for all sequences but *Alternate*, which several subjects segmented into 8 groups of 2 items, while the shortest LoT expression encodes it as a single group of 16 alternating items. **B)** Deviant detection task. Group-averaged sensitivity (d') and response time for each sequence is plotted against LoT complexity. A significant linear relationship with LoT complexity was found in both cases. The Pearson correlation coefficient and associated p-value are reported. Error bars represent one standard error of the mean across participants (SEM).

In summary, using a partially different set of sequences, we replicated the behavioral findings of Planton et al. (2021) showing that, especially for long sequences that largely exceed the storage capacity in working memory, violation detection (an index of learning quality) and response speed (potentially indexing the degree of predictability) were well predicted by our language-of-thought model of sequence compression.

## fMRI data

### A positive effect of complexity during sequence learning and tracking

As predicted, during the habituation phase (i.e. during sequence learning), activation mostly increased with sequence complexity in a broad and bilateral network involving supplementary motor area (SMA), precentral gyrus (preCG) abutting the dorsal part of Brodmann area 44, cerebellum (lobules VI and VIII), superior and middle temporal gyri (STG/MTG), and the anterior intraparietal sulcus region (IPS, close to its junction with the postcentral gyrus) (Figure 3A and Table 1). These regions partially overlapped with those observed in sequence learning for a completely different domain, yet a similar language: the visuo-spatial sequences of Wang et al. (2019). In the opposite direction, a reduction of activation with complexity was seen in a smaller network, mostly corresponding to the default-mode network, which was increasingly deactivated as working memory load increased (Mazoyer et al., 2001; Raichle, 2015): medial frontal cortex, left middle cingulate gyrus, left angular gyrus (AG) and left pars orbitalis of the inferior frontal gyrus (IFGorb) (Table 1).

We then computed the same contrast with the standard trials of the test phase (sequences without violation). The network of areas showing a positive complexity effect was much smaller than during habituation: it included bilateral superior parietal cortex extending into the precuneus, left dorsal premotor area as well as two cerebellar regions (right lobule IV, left lobule VIII) (Figure S1, Table S1). These areas were also found during the habituation phase, although the (predominantly left) parietal superior / precuneus activation was larger and extended more posteriorily than during habituation. These regions may constitute the minimal network needed to track sequences. Regions showing a negative LoT complexity effect in standard trials (reduced activation for increasing complexity) were more numerous: medial frontal regions, middle cingulate gyri, bilateral angular gyrus, bilateral anterior part of the inferior temporal gyrus, bilateral putamen, as well as left frontal orbital region and left occipital gyrus. Here again, they largely resemble what was already observed in habituation trials (i.e. a deactivation of a default mode network), with a few additional elements such as the putamen.

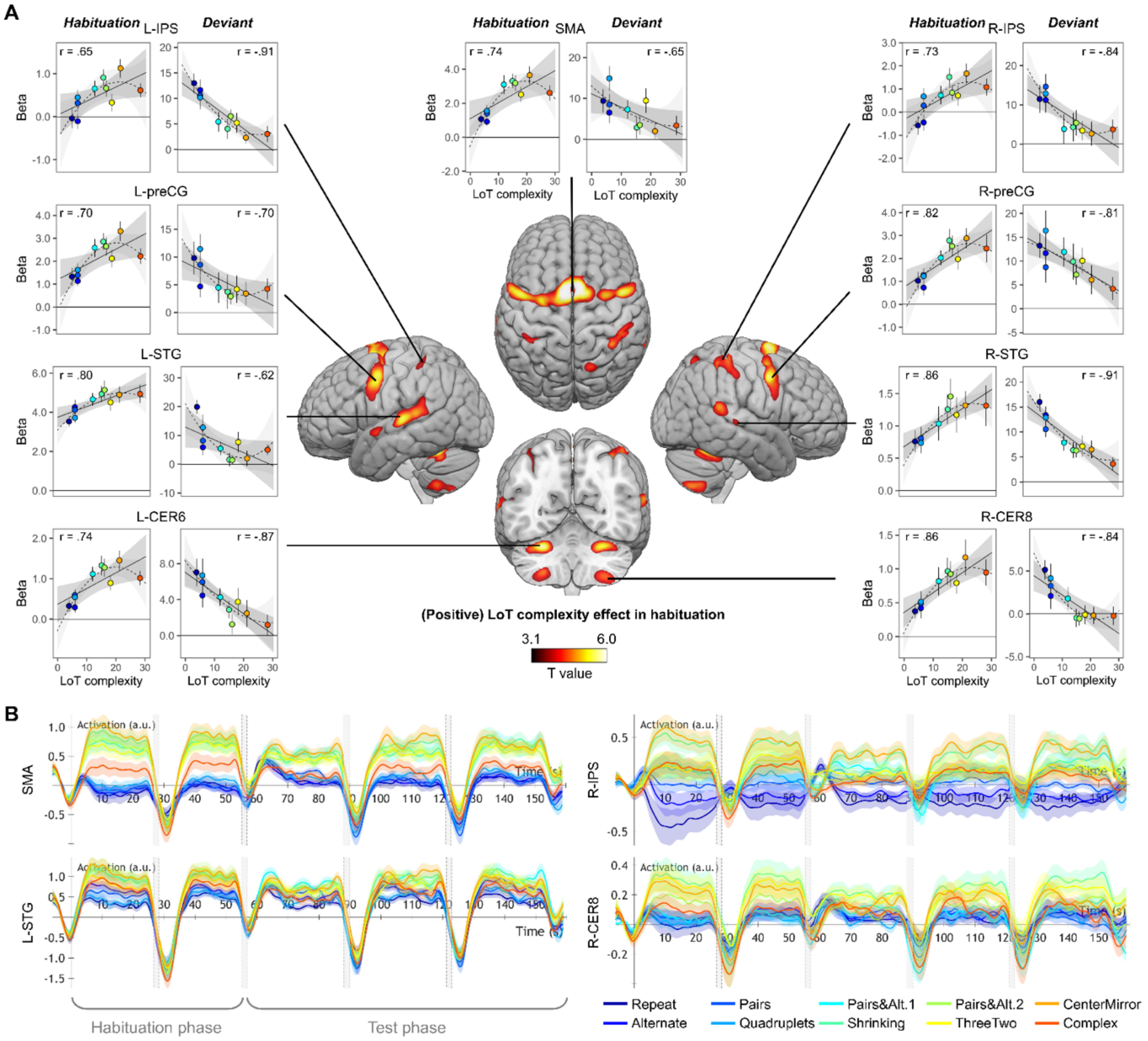
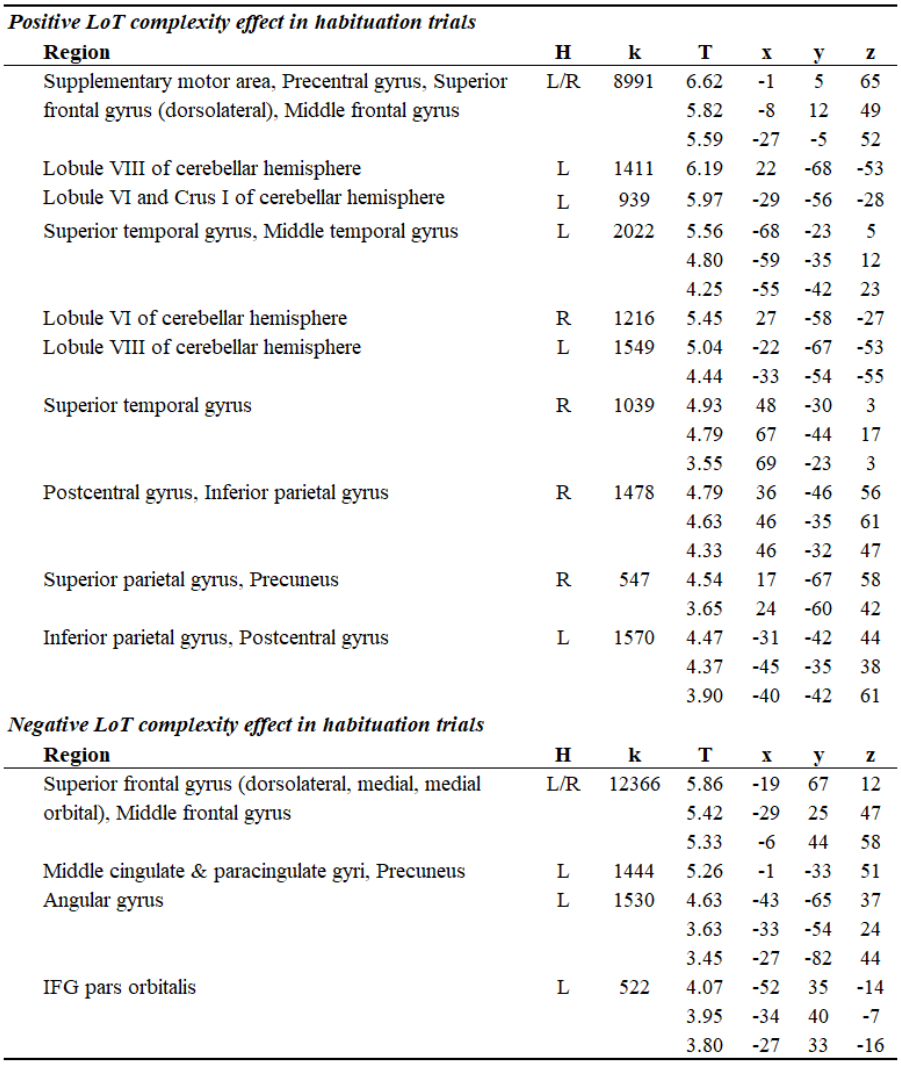


**Figure 3. Sequence complexity in the proposed language of thought (LoT) modulates fMRI responses to standard and deviant sequences. A)** brain areas showing an increase in activation with sequence LoT complexity during habituation (voxel-wise p < .001, uncorrected; cluster-wise p < .05, FDR corrected). Scatterplots represent the group-average activation for each of the ten sequences as a function of their LoT complexity (left panels: habituation trials; right panels, deviant trials) in each of nine ROIs. Data values are from a cross-validated participant-specific ROI analysis. Error bars represent SEM. Linear trend are represented by a solid line (with 95% CI in dark grey) and quadratic trend by a dashed line (with 95% CI in light grey). Pearson linear correlation coefficients are also reported. **B)** Time course of group-averaged BOLD signals for each sequence, for four representative ROIs. Each mini-session lasted 160-seconds and was composed of 5 blocks (2 habituation and 3 tests) interspersed with short rest periods of variable duration (depicted in light gray). The full time course was reconstituted by resynchronizing the data at the onset of each successive block (see Methods). Shading around each time course represents one SEM.

**Table 1. Coordinates of brain areas modulated by LoT complexity during habituation**



| *Positive LoT complexity effect in habituation trials* | | | | | | |
|---|---|---|---|---|---|---|
| **Region** | **H** | **k** | **T** | **x** | **y** | **z** |
| Supplementary motor area, Precentral gyrus, Superior frontal gyrus (dorsolateral), Middle frontal gyrus | L/R | 8991 | 6.62 | -1 | 5 | 65 |
| | | | 5.82 | -8 | 12 | 49 |
| | | | 5.59 | -27 | -5 | 52 |
| Lobule VIII of cerebellar hemisphere | L | 1411 | 6.19 | 22 | -68 | -53 |
| Lobule VI and Crus I of cerebellar hemisphere | L | 939 | 5.97 | -29 | -56 | -28 |
| Superior temporal gyrus, Middle temporal gyrus | L | 2022 | 5.56 | -68 | -23 | 5 |
| | | | 4.80 | -59 | -35 | 12 |
| | | | 4.25 | -55 | -42 | 23 |
| Lobule VI of cerebellar hemisphere | R | 1216 | 5.45 | 27 | -58 | -27 |
| Lobule VIII of cerebellar hemisphere | L | 1549 | 5.04 | -22 | -67 | -53 |
| | | | 4.44 | -33 | -54 | -55 |
| Superior temporal gyrus | R | 1039 | 4.93 | 48 | -30 | 3 |
| | | | 4.79 | 67 | -44 | 17 |
| | | | 3.55 | 69 | -23 | 3 |
| Postcentral gyrus, Inferior parietal gyrus | R | 1478 | 4.79 | 36 | -46 | 56 |
| | | | 4.63 | 46 | -35 | 61 |
| | | | 4.33 | 46 | -32 | 47 |
| Superior parietal gyrus, Precuneus | R | 547 | 4.54 | 17 | -67 | 58 |
| | | | 3.65 | 24 | -60 | 42 |
| Inferior parietal gyrus, Postcentral gyrus | L | 1570 | 4.47 | -31 | -42 | 44 |
| | | | 4.37 | -45 | -35 | 38 |
| | | | 3.90 | -40 | -42 | 61 |
| *Negative LoT complexity effect in habituation trials* | | | | | | |
| **Region** | **H** | **k** | **T** | **x** | **y** | **z** |
| Superior frontal gyrus (dorsolateral, medial, medial orbital), Middle frontal gyrus | L/R | 12366 | 5.86 | -19 | 67 | 12 |
| | | | 5.42 | -29 | 25 | 47 |
| | | | 5.33 | -6 | 44 | 58 |
| Middle cingulate & paracingulate gyri, Precuneus | L | 1444 | 5.26 | -1 | -33 | 51 |
| Angular gyrus | L | 1530 | 4.63 | -43 | -65 | 37 |
| | | | 3.63 | -33 | -54 | 24 |
| | | | 3.45 | -27 | -82 | 44 |
| IFG pars orbitalis | L | 522 | 4.07 | -52 | 35 | -14 |
| | | | 3.95 | -34 | 40 | -7 |
| | | | 3.80 | -27 | 33 | -16 |

## A negative effect of complexity on deviant responses

The effect of LoT complexity at the whole brain level was first assessed on the responses to all deviant stimuli (whether detected or not). A positive linear effect of complexity was only found in a small cluster of the medial part of the superior frontal gyrus (SFG) (Table S2). As predicted, a much larger network showed a negative effect (i.e. reduced activation with complexity or increased activation for less complex sequences): bilateral postcentral gyrus (with major peak in the ventral part), supramarginal gyrus (SMG), IPS, STG, posterior MTG, ventral preCG, Insula, SMA and middle cingulate gyrus, cerebellum (lobules VI, VIII, and vermis) (red activation map of Figure 4, Table S2). This network is thus the possible brain counterpart of the increase in deviant detection performance observed as sequences become less and less complex. However, this result could be partly due to a motor effect, since manual motor responses to deviants were less frequent for more complex sequences, as attested by an effect of LoT complexity on sensitivity. We therefore computed the same contrast using an alternative GLM modeling only deviant trials to which the participant correctly responded (note that this model consequently included fewer trials, especially for higher complexity sequences). Negative effects of LoT complexity were still present in this alternative model, now unconfounded by motor responses. As shown in Figure 4 (yellow) the negative effect network was a subpart of the network identified in the previous model, and concerned bilateral STG, MTG, SMG/postcentral gyrus, Insula, SMA and middle cingulate gyrus. A positive effect was still present in a medial SFG cluster, part of the default-mode network showing less deactivation for deviants as complexity increased.

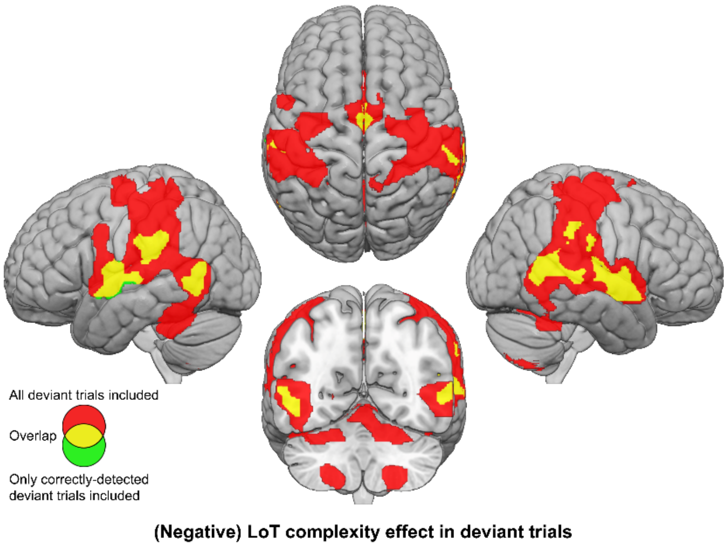

All deviant trials included
Overlap
Only correctly-detected deviant trials included
(Negative) LoT complexity effect in deviant trials

**Figure 4. Brain responses to deviants decrease with LoT complexity.** Colors indicate the brain areas whose activation on deviant trials decreased significantly with complexity, in two distinct general linear models (GLMs): one in which all deviant stimuli were modeled (red), and one in which only correctly-detected deviant stimuli were modeled (green) (voxel-wise p < .001, uncorrected; cluster-wise p < .05, FDR corrected). Overlap is shown in yellow.

### ROI analyses of the shape of the complexity effect

We next used individual ROIs to measure the precise shape of the complexity effect and test the hypothesis that (1) activation increases with complexity but may reach a plateau or decrease for the most complex, incompressible sequence; and (2) on deviant trials, the complexity effect occurs in the opposite direction. We designed cross-validated individual ROI analyses, which consisted in (1) using half of the runs to identify responsive individual voxels within each ROI, using a contrast of positive effect of complexity during habituation; and (2) using the other half to extract the activation levels for each standard or deviant sequence. We focused on nine areas that exhibited a positive complexity contrast in habituation (figure 3), where the effect was robust and was computed on the learning phase of the experiment, therefore uncontaminated by deviant stimuli and manual motor responses.

In each ROI, mixed effect models with participants as the random factor were used to assess the replicability of the linear effect of complexity during habituation. A significant effect was found in all ROIs (after Bonferroni correction for nine ROIs), although with variable effect size: SMA: β estimate = 0.10, t(21) = 5.37, p.corr < .0003; L-STG: β = 0.06, t(21) = 4.75, p.corr < .001; L-CER6: β = 0.04, t(21) = 4.73, p.corr < .002; R-IPS: β = 0.07, t(21) = 4.08, p.corr < .005; L-preCG: β = 0.07, t(21) = 3.98, p.corr < .007; R-preCG: β = 0.08, t(21) = 3.8, p.corr < .01; R-STG: β = 0.03, t(21) = 3.35, p.corr < .03; R-CER8: β = 0.03, t(21) = 3.32, p.corr < .03 and L-IPS: β = 0.03, t(21) = 3.25, p.corr < .04. These results are illustrated in Figure 4A, showing the linear regression trend with values averaged per condition across participants. The addition of a quadratic term was significant for seven ROIs (SMA, L-CER6, L-IPS, L-preCG, L-STG, R-CER8 and R-IPS), but did not reached significance in R-preCG nor in R-STG. This effect was always negative, indicating that the activation increase with complexity reached saturation or decreased from the most complex sequence (see dashed lines in the scatter plots of Figure 4A).

We also examined the time course of activation profiles within each mini-session of the experiment, i.e. two habituation blocks followed by three test blocks. As shown in Figure 4B (see Figure S2 for all 9 ROIS), the activation time courses showed a brief activation to sequences, presumably corresponding to a brief search period. 5 to 10 seconds following the first block onset, however, activation quickly dropped to a similar and very low activation, or even a deactivation below the rest level, selectively for the 4 lowest-complexity sequences which involved only simple processes of transition probabilities or chunking. For other sequences, the BOLD effect shot up in rough proportion to complexity, yet with a midlevel amplitude for the most complex sequence reflecting the saturation, quadratic effect noted earlier. Thus, in 5-10 seconds, the profile of the complexity effect was firmly established, and it remained sustained over time during habituation and, with reduced amplitude, during test blocks. This finding indicated that the same areas were responsible for discovering the sequence profile and for monitoring it for violations during the test period. The profile was similar across regions, with one exception: while most areas showed the same, low activation to the first four, simplest sequences, the left and right IPS showed an increasing activation as a function of the number of items in a chunk (ABABAB… = 1; AABBAA… = 2; AAAABBBB… = 4).

This observation fits with the hypothesis that these regions are involved in numerosity representation, and may therefore implement the "for loops" postulated in our language.

The ROI analyses were next performed with data from the deviant trials, in order to test whether areas previously identified as sensitive to sequence complexity when learning the sequence also showed an opposite modulation of their response to deviant trials. All ROIs indeed showed a significant negative effect of LoT complexity: R-STG: β = -0.46, t(197) = -8.01, p.corr < .0001; L-IPS: β = -0.44, t(197) = -6.35, p.corr < .0001; R-CER8: β = -0.23, t(197) = -5.09, p.corr < .0001; L-STG: β = -0.45, t(197) = -4.61, p.corr < .0001; L-CER6: β = -0.23, t(197) = -4.57, p.corr < .0001; R-preCG: β = -0.37, t(197) = -3.92, p.corr < .002; R-IPS: β = -0.49, t(197) = -3.85, p.corr < .002; SMA: β = -0.33, t(197) = -3.57, p.corr < .004 and L-preCG: β = -0.27, t(197) = -2.9, p.corr < .04. Interestingly, unlike during habituation, the addition of a quadratic term did not improve the regression except in a single area, L-STG: β = 0.05, t(196) = 3.9, p.corr < .002. Smaller effects of the quadratic term were present in three other areas, but they were not significant after Bonferonni correction: R-CER8: β = 0.02, t(196) = 2.68, p < .009; R-STG: β = 0.02, t(196) = 2.42, p < .02 and L-IPS: β = 0.02, t(196) = 2.3, p < .03.

As in the whole-brain analysis, we finally conducted a complementary analysis using activation computed with correctly-detected deviants trials only. The linear LoT complexity was now only significant in four of the nine ROIs: R-STG: β = -0.48, t(197) = -7.64, p.corr < .0001; L-IPS: β = -0.34, t(197) = -4.54, p.corr < .0001; L-STG: β = -0.42, t(197) = -4.12, p.corr < .0006; R-CER8: β = -0.15, t(197) = -3.17, p.corr < .02. When adding a quadratic term, no significant effects were observed at the predefined threshold, although uncorrected ones were present for L-STG: β = 0.03, t(196) = 2.37, p < .02 and R-CER8: β = 0.01, t(196) = 2.05, p < .05.

### Overlap with the brain networks for language and mathematics

Past and present behavioral results suggest that an inner "language" is required to explain human working memory for auditory sequences – but is this language similar to natural language? Or to the language of mathematics, and more specifically geometry, from which it is derived (Al Roumi et al., 2021; Amalric et al., 2017; Wang et al., 2019) ? By including in our fMRI protocol an independent language and mathematics localizer experiment, we tested whether the very same cortical sites are involved in natural sentence processing, mathematical processing, and auditory sequences.

At the whole-brain group level, large amount of overlap was found between the mathematics network (whole-brain mental computation > sentences processing contrast, in a $2^{nd}$ level ANOVA analysis of the localizer experiment) and the LoT complexity network (see Figure 5A): SMA, bilateral precentral cortex, bilateral anterior IPS, and bilateral cerebellum (lobules VI). Some overlap was also present, to a lower extent, with the language network (auditory and visual sentences > auditory and visual control stimuli) and the LoT complexity network: left STG, SMA, left precentral gyrus, and right cerebellum.

Such group-level overlap, however, could be misleading since they involve a significant degree of intersubject smoothing and averaging. For a more precise assessment of overlap, we extracted, for each subject and within each of 7 language-related and 7 math-related ROIs (see figure 5), the subject-specific voxels that responded, respectively, to sentence processing and to mental calculation (same contrasts as above, but now within each subject). We then extracted the results from those ROIs and examined their variation with LoT complexity in the main experiment (during habituation). In the language network, a significant positive effect of LoT complexity during the habituation phase was only found in left IFGoper: $\beta = 0.03$, $t(197) = 4.25$, p.corr < .0005 (Figure 5B). In fact, most other language areas showed either no activation or were deactivated (e.g. IFGorb, aSTS, TP, TPJ). As concerns deviants, a significant negative effect of LoT complexity was found in left IFGoper: $\beta = -0.23$, $t(197) = -3.04$, p.corr < .04; and in left pSTS: $\beta = -0.24$, $t(197) = -3.27$, p.corr < .02. The quadratic term was never found significant.

On the contrary, in the mathematics-related network, all areas showed a positive LoT complexity effect in habituation (Figure 5B): SMA: $\beta = 0.05$, $t(197) = 5.6$, p.corr < .0001; left preCG/IFG: $\beta = 0.05$, $t(197) = 5.03$, p.corr < .0001; right IPS: $\beta = 0.05$, $t(197) = 4.69$, p.corr < .0001; right preCG/IFG: $\beta = 0.05$, $t(197) = 4.56$, p.corr < .0002; right SFG: $\beta = 0.04$, $t(197) = 4$, p.corr < .002; left IPS: $\beta = 0.04$, $t(197) = 3.78$, p.corr < .003 and left SFG: $\beta = 0.02$, $t(197) = 3.15$, p.corr < .03. The quadratic term in the second model was also significant for three of them: SMA: $\beta = -0.01$, $t(196) = -4.11$, p.corr < .0009; right preCG/IFG: $\beta = 0$, $t(196) = -3.21$, p.corr < .03 and left preCG/IFG: $\beta = 0$, $t(196) = -3.1$, p.corr < .04. A negative complexity effect for deviant trials reached significance in four areas: left IPS: $\beta = -0.42$, $t(197) = -4.44$, p.corr < .0003; left preCG/IFG: $\beta = -0.48$, $t(197) = -4.31$, p.corr < .0004; right IPS: $\beta = -0.41$, $t(197) = -4$, p.corr < .002 and SMA: $\beta = -0.29$, $t(197) = -2.97$, p.corr < .05. Their response pattern was not significantly quadratic.

To summarize, all dorsal regions previously identified as involved in mathematical processing regions were sensitive to the complexity of our auditory binary sequences, as manifested by an increase, up to a certain level of complexity, during habituation; and, for most regions, a reduction of the novelty to deviants (especially for SMA, left preCG and IPS). Such a sensitivity to complexity was conspicuously absent from language areas, except for the left pars opercularis of the IFG.

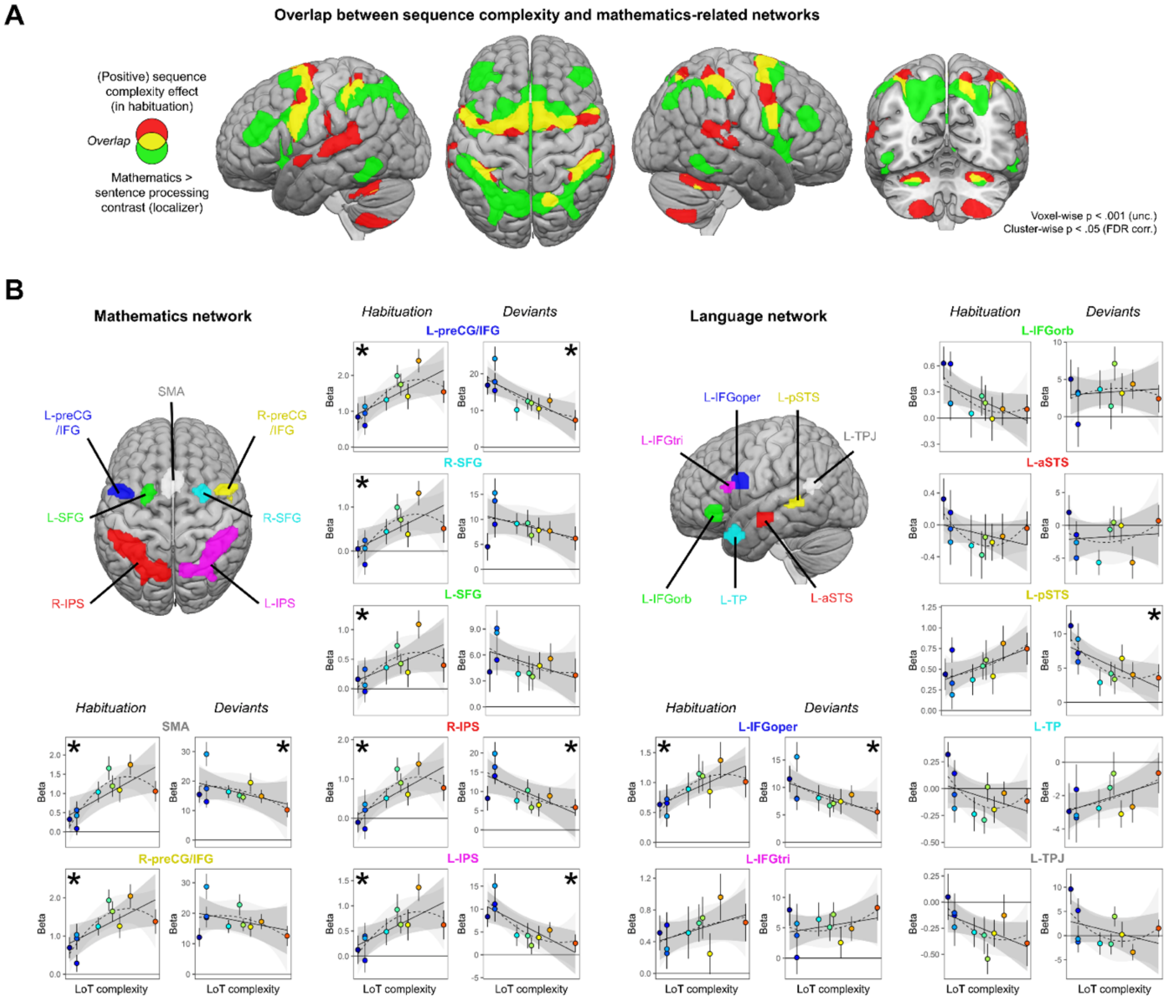


**Figure 5. Sequence complexity effects in mathematics and language networks. A)** Overlap between the brain areas showing an increase of activation with sequence LoT complexity during habituation in the main experiment (in red) and the brain areas showing an increased activation for mathematical processing (relative to simple listening/reading of non-mathematical sentences) in the localizer experiment (in green; both maps thresholded at voxel-wise p < .001 uncorrected, cluster-wise p < .05, FDR corrected). Overlap between the two activation maps is shown in yellow. **B)** Overview of the 7 search volumes representing the mathematics network (left) and the 7 search volumes representing the language network (right) used in the ROI analyses. Within each ROI, each scatter plot represents the group-average activation for each of the ten sequences according to their LoT complexity, for habitation blocks and for deviant trials (same format as figure 3). A star (*)indicates significance of the linear effect of LoT complexity in a linear mixed-effects model.

## MEG Results

The temporal resolution of fMRI did not permit tracking the successive sequence items, but only the average activity they induced. This lack of temporal resolution may have prevented us from detecting subtle effects, particularly in the timing of responses to deviants. To address this concern, a similar paradigm was tested with MEG. To maximize signal-to-noise, especially on the rare deviant trials, only seven sequences were selected (figures 1 and 6). Unlike the fMRI experiment, during MEG we merely asked participants to listen carefully to the presented sequences of sounds, without providing any button response, thus yielding pure measures of violation detection uncontaminated by the need to respond.

### Neural signatures of complexity at the univariate level

We first determine if a summary measure of brain activity, the Global Field Power, is modulated by sequence complexity. To do so, we consider the brain responses to sounds occurring in the *habituation* phase, to non-deviant sounds occurring in the test phase (referred to as *standard* sounds) and to *deviant* sounds. On *habituation* trials, the late part of the auditory response (108ms – 208ms) correlated positively with complexity (p = 0.00024, see shaded area in the top panel of Figure 6A): the more complex the sequence, the larger the brain response. On *standard* trials, this modulation of the GFP by complexity had vanished (middle panel of Figure 6A). Finally, as predicted, the GFP computed on the *deviant* exhibited the reversed effect, i.e. a negative correlation with complexity on the 116 - 300 ms time-window (p = 0.0005) and on the 312 – 560 ms time-window (p = 0.0005), indicating that *deviants* elicit larger brain responses in sequences with lower complexity (bottom panel of Figure 6A).

To better characterize the mechanisms of sequence coding, we ran a linear regression of the evoked responses to sounds as a function of sequence complexity. Regression coefficients of the sequence complexity predictor were projected to source space. The results showed that complexity effects were present in temporal and precentral of the cortex. To assess the significance of the regression coefficients, we ran a spatiotemporal cluster-based permutation test at the sensor level. Several significant clusters were found for each of the 3 trial types (*habituation :* cluster 1 from 72 to 216 ms, p = 0.0004, cluster 2 from 96 to 212 ms, p = 0.0002; *standard* : cluster 1 from 96 to 180ms,  p = 0.0038, cluster 2 from 96 to 184 ms, p

= 0.001; *deviant* : cluster 1 from 60 to 600 ms, p = 0.0002, cluster 2 from 56 to 600 ms, p = 0.0002). Figure 6 illustrates one significant cluster for each trial type.

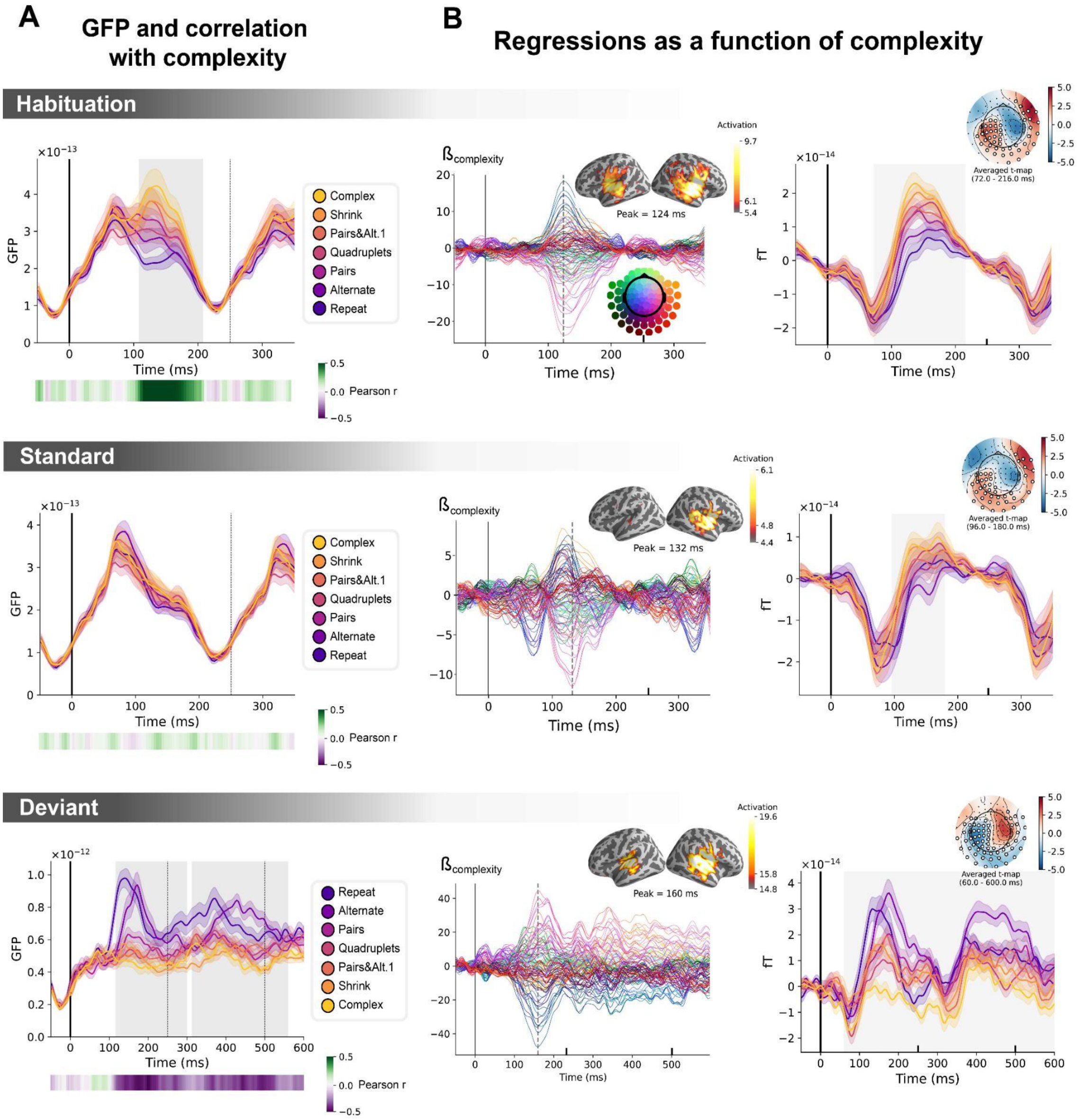


**Figure 6. Sequence complexity in the proposed language of thought (LoT) modulates MEG signals to habituation, standard and deviant trials. A)** Global field power computed for each sequence (see color legend) from the evoked potentials of the *Habituation, Standard* and *Deviant* trials*.* 0 ms indicates sound onset. Note that the time-window ranges until 350 ms for *Habituation* and *Standard* trials (with a new sound onset at S0A=250 ms), and until 600 ms for *Deviant* trials and for the others. Significant correlation with sequence complexity was found in *Habituation* and *Deviant* GFPs and are indicated by the shaded areas. **B)** Regressions of MEG signals as a function of sequence complexity. Left: amplitude of the regression coefficients ß of the complexity regressor for each MEG sensor. Insets show the projection of those coefficients in source space at the maximal amplitude peak, indicated by a vertical dotted line. Right: spatiotemporal clusters where regression coefficients were significantly different from 0. While several clusters were found (see Text and Figure S3), for the sake of illustration, only one is shown for each trial type. The clusters involved the same sensors but on different time windows (indicated by the shaded areas) and with an opposite T-value for *Deviant* trials. Neural signals were averaged over significant sensors for each sequence type and were plotted separately.

The clusters shown involve the same sensors but exhibit opposite regression signs for the brain responses to *Deviant* sounds, suggesting that, as in fMRI, the same brain regions are involved in the processing of standard and deviant items but are affected by complexity in an opposite manner.

## Controlling for local transition probabilities

Several studies have shown that the human brain is sensitive to the statistics of sounds and sound transitions in a sequence (Maheu et al., 2019; Meyniel et al., 2016; Näätänen et al., 1989; Todorovic et al., 2011; Todorovic & Lange, 2012; Wacongne et al., 2012), including in infants 15/10/2022 17:30:00(Saffran et al., 1996). When listening to probabilistic binary sequences of sounds, early brain responses reflect simple statistics such as item frequency while later brain responses reflect more complex, longer-term inferences (Maheu et al., 2019). Since local surprise and global complexity were partially correlated in our sequences, could surprise alone account for our results? To disentangle the contributions of transition probabilities and sequence structure in the present brain responses, we regressed the brain signals as a function of complexity and of surprise based on transition probabilities. To capture the latter, we added several predictors: the presence of a repetition or an alternation and the surprise of an ideal observer that makes optimal inferences about transition probabilities from the past 100 items (see Maheu et al., 2019 for details). Both predictors were computed for two consecutive items: the one at stimulus onset (t=0ms) and the next item (t=250ms later) and included together with LoT complexity as multiple regressors of every time point.

Figure S4 shows the temporal profile of the regression coefficient for sequence complexity for each MEG sensor and its projection onto the source space, once these controlling variables were introduced. The contribution of auditory regions was slightly diminished compared to the simple regression of brain signals as a function of complexity (Figure S3). To assess the significance of the regression coefficient, we ran a spatiotemporal cluster-based permutation test at the sensor level. Several significant clusters were found for each of the 3 trial types (*habituation :* cluster 1 from 96 to 244 ms, p = 0.0162, cluster 2 from 112 to 220 ms, p = 0.014 ; *standard* : cluster 1 from 104.0 to 180.0 ms, cluster-value= 1.50, p = 0.0226, cluster 2 from 100 to 220 ms, p = 0.0004; *deviant* : cluster 1 from 224 to 600 ms, p = 0.0088, cluster 2 from 116 to 600 ms, p = 0.0006; see Figures S3 and S4 for complete cluster profiles). The results remained even when the surprise regressors were entered first, and then the regression on complexity was performed on the residuals (figure S4, right column). In summary, the positive effect of complexity on habituation and standard trials, and its negative effect on deviant trials, were not solely due to local transition-based surprise signals.

### Time-resolved decoding of violation responses

The above results were obtained by averaging sensor data across successive stimuli and across participants. A potentially more sensitive analysis method is multivariate decoding (King & Dehaene, 2014), which searches, at each time point and within each participant, for an optimal pattern of sensor activity reflecting a given type of mental representation. Therefore, to further characterize the brain representations of sequence structure and complexity, we next used multivariate time-resolved analyses, which allowed us to track sequence coding for each item in the sequence, at the millisecond scale.

We trained a decoder to classify all standard versus all deviant trials (El Karoui et al., 2015; King et al., 2013). As the two versions of the same sequence were presented on two separated runs (respectively starting with sound ‘x’ or ‘Y’), we trained and tested the decoder in a leave-one-run out manner, thus forcing it to identify non-stimulus specific sequence violation responses. In addition, and most importantly, we selected standard trials that matched the deviants’ ordinal position, which was specific to each sequence (see figure 1, orange lines). Figure 7 shows the average projection on the decision vector of the classifier’s predictions on left-out data for the different sequences, when tested on both position-matched *Deviants* versus *Standards* (Figure 7A) and on *Habituation* trials (Figure 7B). Significance was determined by temporal cluster-based permutation tests.

Decoding of deviants reached significance for all sequences except for the most complex one (with only a short burst of significance for the 2nd most complex *Shrink* sequence). For the simplest *Repeat* and *Alternate* sequences, which could be learned solely based on transition probabilities, a sharp initial mismatch response was seen, peaking at ~150 ms. For all other sequences, the decoder exhibited a later, slower, lower-amplitude and sustained development of above-chance performance, suggesting that deviant items elicit decodable long-lasting brain signals. A temporal cluster-based permutation test on Pearson correlation with sequence complexity showed that the decoding of violations strongly correlated with complexity over a broad time-window (~90 - 580 ms).

The time courses of the decoder performance on habituation trials also revealed a clear hierarchy in the time it took for the brain to decide that a given tone was not a deviant (figure 7B). The seven curves were ordered by predicted sequence complexity. Thus, the decoder's classification as Standard, quantified as the projection on the decision vector, decreased significantly with sequence complexity over two time windows (~90 - 220 ms and ~330 - 460 ms). This suggests that the more the sequence is complex, the more brittle its classification as Standard is.

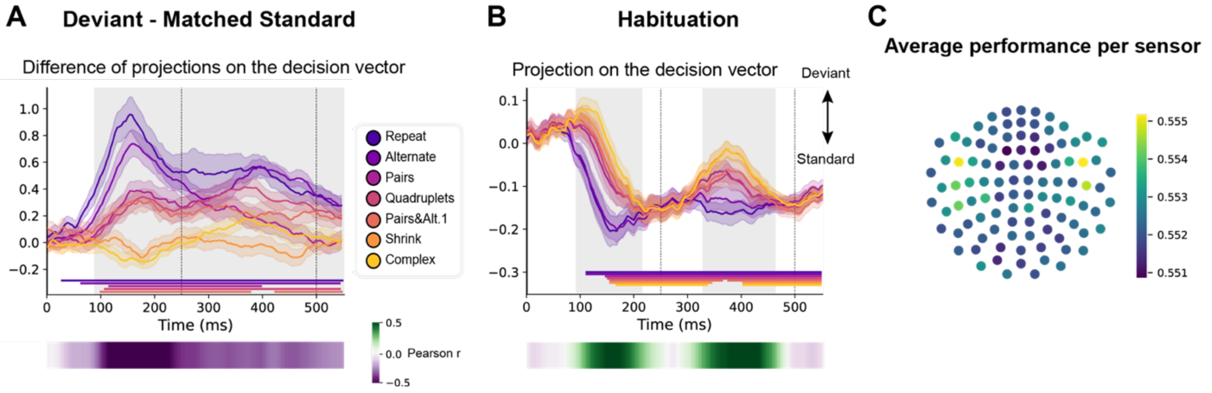


**Figure 7. Multivariate decoding of deviant trials from MEG signals, and its variation with sequence complexity. A,** A decoder was trained to classify standard from deviant trials from MEG signals at a given time point. We here show the difference in the projection on the decision vector for *Standard* and *Deviant* trials, that is a measure of the decoder's accuracy. The decoder was trained jointly on all sequences, but its performance is plotted here for left-out trials separately for each sequence type. Shaded areas indicate s.e.m. and colored lines at bottom indicate significant time windows (p<0.05 corrected) obtained from cluster-based permutation test on the full window. The heatmap at the bottom represents the correlation of the performance with sequence complexity (Pearson's r). Correlation is significant in the gray shaded time-window in the main graph (two tailed p<0.05, temporal cluster-based permutation test). **B,** Projection on the decision vector for *Habituation* trials. The early brain response is classified as deviant but later as standard. This projection time-course is increasingly delayed as a function of sequence complexity (same format as **A**). **C,** sensor map showing the relative contribution of each sensor to overall decoding performance. At the time of maximal overall decoding performance (165 ms) we trained and tested 4000 new decoders that used only a subset of 40 gradiometers at 20 sensor locations. For each sensor location, the color on the maps in the right column indicates the average decoding performance when this sensor location was used in decoding, thus assessing its contribution to overall decoding.

## Decoder performance over the full extent of each sequence

To characterize the time course of brain activity over the entire course of each sequence, we projected each MEG time point onto the decoding axis of the standard/deviant decoder trained on data from a 130-210 ms time window (Figure 8). The projection was computed separately for each sequence, separately for habituation, standard, and the four possible positions of deviant trials. We determined if deviants differed from standards using a cluster-based permutation test on a 0 - 600ms window after each violation (colored lines at the bottom of each sequence in Figure 8A).

All individual deviants elicited a significant decodable response except for the two highest-complexity sequences: *Shrinking* and *Complex* (failure at all positions exception the last one: 15). Interestingly, for the alternate sequence, two consecutive peaks indicate that, when a single repetition is introduced in an alternating sequence (e.g. ABAB**B**BAB… instead of ABABABAB…), the brain interprets it as two consecutive violations, probably due to transition probabilities, as each of the B items is predicted to be followed by an A.

Most crucially, the analysis of specific violation responses allowed us to evaluate the range of properties that humans use to encode sequences, and to test the hypothesis that they integrate numerical and structural information at multiple nested levels (Wang et al., 2015). First, within a chunk of consecutive items, they detect violations consisting in both chunk shortening (1 repeated tone instead of 2 in *Pairs*; 3 tones instead of 4 in *Quadruplets*) and chunk lengthening (3 repeated tones instead of 2, as well as 5 instead of 4). The contrast between those two sequences clearly shows that participants possess a sophisticated context-dependent representation of each sequence. Thus, their brain emits a violation response upon hearing 3 consecutive items (AA**A**) within the *Pairs* sequence, where it is unexpected, but not when the same sequence occurs within the *Quadruplets* sequence. Conversely, participants are surprised to hear the transition BBAA**B** in the *Quadruplet* context, but not in the *Pairs* context. Finally, in the *Pairs+Alt.1* sequence, such context dependence changes over time, thus indicating an additional level of nesting: at positions 9-12, subjects expect to hear two pairs (AABB) and are surprised to hear A**B**BB (unexpected alternation), but just a second later, at positions 13-16, they expect an alternation (ABAB) and are surprised to hear A**A**AB (unexpected repetition). Similar, though less significant, evidence for syntax-based violation responses are present in the *Shrinking* sequence, which also ends with two pairs and an alternation.

Figure 8 also shows in great detail how the participants' brain fluctuates between predictability (in blue) and violation detection (in red) during all phases of the experiment. Initially, during habituation (top line), sequences are partially unpredictable, as shown by red responses to successive stimuli, but that effect is strongly modulated by complexity, as previously reported (red responses, particularly for the most complex sequences). In a sense, while the sequence is being learned, all items in those sequences appear as deviants. As expected, after habituation, the deviancy response to standards is much reduced, but still ordered by complexity. Higher-complexity sequences such as *Shrinking* thus creates a globally less predictable environment (red colors) relative to which the violation responses to deviants appear to be reduced.

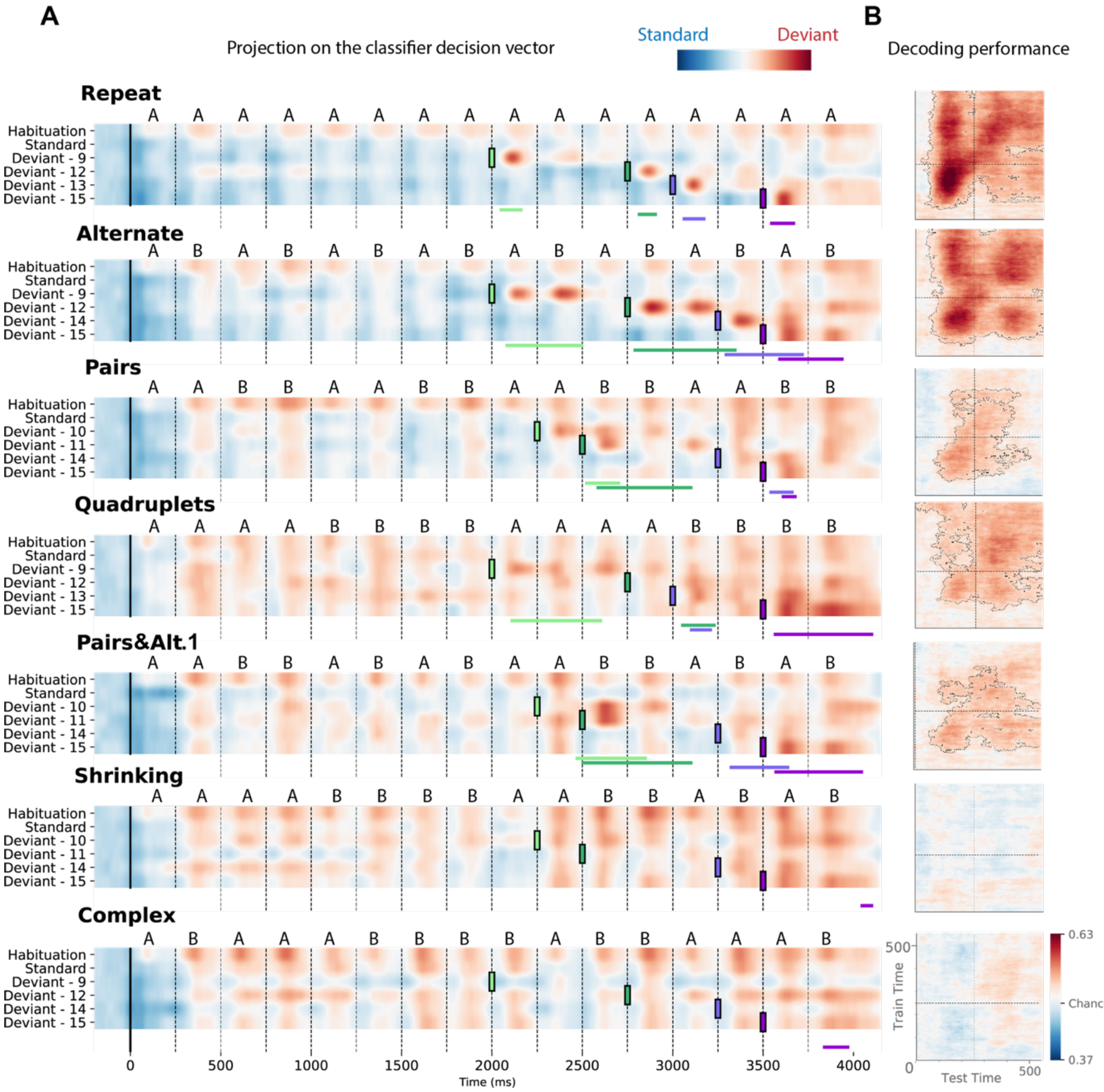


**Figure 8. Time course of the deviancy decoder across the different types of sequences and deviants. A**) Average projection of MEG signals onto the decoding axis of the standard/deviant decoder. For each sequence, the time course of the projection was computed separately for habituation trials, standard trials, and for the four types of trials containing a deviant at a given position. The figure shows the average output of decoders trained between 130 ms and 210 ms post-deviant. Red indicates that a trial tends to be classified as a deviant, blue as a standard. Colored lines at the bottom of each graph indicate time windows with a significant deviant signal (cluster permutation test comparing deviants and standards in a 0-600 ms window after deviant onset). **B**) Average generalization-across-time (GAT) matrix showing decoding performance 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 (see Methods). Simpler sequences exhibit overall greater performance as well as larger time windows of significance. We note that, while deviancy detection does not reach significance for Shrinking and Complex sequences in the GAT matrices, violation signals reached significance for deviant position 15.

Figure 8B also shows how the Standard-Deviant decoder generalizes over time, separately for each sequence. The performance for the *Repeat* sequence exhibited a peak corresponding to the deviant item's presentation (~ 150 ms) and a large and a partial square pattern, indicating a sustained maintenance of the deviance information. The performance for the *Alternate* sequence shows 4 peaks spaced by the SOA, corresponding to the two deviant transitions elicited by the deviant item. *Pairs, Quadruplets* and *Pairs+Alt.1* sequences still show significant decoding times but not *Shrinking* and *Complex* sequences, indicating that the ability to decode deviant signals decreases with complexity.

## Discussion

The goal of this study was to characterize the mental representation that humans utilize to encode binary sequences of sounds in working memory and detect occasional deviants. The results indicate that, in the human brain, deviant responses go way beyond the sole detection of violations in habitual sounds (May & Tiitinen, 2010) or in transition probabilities (Wacongne et al., 2012), and are also sensitive to more complex, larger-scale regularities (Bekinschtein et al., 2009; Bendixen et al., 2007; Maheu et al., 2019; Schröger et al., 2007; Wacongne et al., 2011; Wang et al., 2015). Instead of merely storing each successive sound in a distinct memory slot (Baddeley, 2003; Baddeley & Hitch, 1974; Botvinick & Watanabe, 2007; Hurlstone et al., 2014), behavioral and brain imaging results suggest that participants mentally compressed these sequences using an algorithmic-like description where sequence regularities (Dehaene et al., 2015) are expressed in terms of combination of simple rules that are recursively integrated (Al Roumi et al., 2021; Planton et al., 2021). Consistently with the predictions of this formal language of thought (LoT), behavioral performance and brain responses were modulated by the minimal description length (MDL) of the sequence, which we term LoT complexity. We discuss those points in turn.

Our behavioral results, obtained during fMRI, fully replicated our previously behavioral work (Planton et al., 2021) showing that, for long sequences, sequence learning difficulty is strongly modulated by minimal description length in our formal language. In absence of any regularity, a 16-item sequence should be way above the normal working memory span. When a deviant was correctly detected, the response time was modulated significantly as a function of LoT complexity, suggesting that novelty detection mechanisms were impacted by sequence structure. Finally, after the experiment, participants were asked to segment with brackets the sequences. The proposed segmentations matched on average the LoT sequence descriptions: participants did not rely solely on the presence of repetitions to segment the sequences, but also relied on transitions between higher level chunks were often identified. For instance, they segmented the *Pairs+Alt.1* sequence as [[AA][BB]][[ABAB]]. These behavioral results confirm that the postulated LoT provides a plausible description of how binary sequences are encoded. They fit with a long line of cognitive psychological research searching for computer-like languages that may capture the human notion of regularity for sequences (Leeuwenberg, 1969; Restle, 1970; Restle & Brown, 1970; Simon, 1972; Simon & Kotovsky, 1963). While the present behavioral evidence is limited, Planton et al. (2021) provided a formal, statistical comparison demonstrating the superiority of LoT complexity against many competing measures such as transition probability, chunk complexity, entropy, subsymmetries, Lempel-Ziv compression, change complexity or algorithmic complexity. In the next sections, we discuss how brain imaging results provide additional information on how sequence compression is implemented in the human brain.

According to our hypothesis, the more complex the sequence, the longer the internal model and the larger the effort to parse it, encode it and maintain it in working memory. Consequently, we expected during the habituation phase larger brain activations for more complex sequences in regions that are involved in auditory sequence encoding. Both fMRI and MEG results support this hypothesis. Importantly, contrary to the fMRI experiment, the MEG did not require overt responses, yet several neural markers, such as Global Field Power, showed a significant increase with sequence complexity (Figure 6A). Furthermore, linear regressions showed that brain activity increased with sequence complexity for a given cluster of electrodes that corresponded to the auditory and inferior frontal regions (Figure 6B).

Many levels of sequence processing mechanisms coexist in the human brain (Dehaene et al., 2015). At a minimum, one should distinguish the coding of transition probabilities between consecutive sounds and of sequence structure as described by the postulated language of thought (Bekinschtein et al., 2009; Maheu et al., 2019; Wacongne et al., 2011). To separate them, we ran a linear multilinear regression model with regressors for transition-based statistics (lower-order statistical properties were not relevant as the overall item frequency was equalized). Even after adding four additional regressions for immediate and longer-term transition statistics, the regressor for complexity was still significant over similar sensor clusters and time-windows (figure S3). As shown in Figure S5, repetition/alternation impacted on both an early peak at 80ms and a later one at 176ms after stim onset, perhaps reflecting sensory bottom-up versus top-down processes. Transition-based surprise exhibited only one peak at 104 ms after stim onset. The 20ms delay between the peaks supports the idea that the first reflects low-level neural adaptation while the second corresponds to a violation of expectations based on transition probabilities. Complexity effects, however, showed a later and more sustained response, extending way beyond 200 ms, in agreement with a distinct rule-based process.

Previous fMRI results led us to expect several prefrontal regions to exhibit an increasing activity with sequence complexity (Badre, 2008; Badre et al., 2010; Barascud et al., 2016; Koechlin et al., 2003; Koechlin & Jubault, 2006; Wang et al., 2019), but no such activation was observed in MEG source reconstruction. This negative result has several potential explanations. First of all, sequence complexity may act as a context effect and therefore may be sustained across time (Barascud et al., 2016; Southwell & Chait, 2018). As we baselined the data on a short time-window before each sound onset, such a constant effect may be removed. Furthermore, frontal brain regions may be too distributed, intermixed and/or too far from the MEG helmet to be faithfully reconstructed. Finally, the fMRI experiment allowed us to clearly identify a large network of brain areas involved in complexity, but recruiting a rather posterior region of prefrontal cortex, the preCG (or dorsal premotor cortex, PMd, bordering on the dorsal part of Brodmann area 44) together with the STG, SMA, cerebellum, and IPS that all exhibited the predicted increase in activity with LoT complexity. All these regions showed the predicted increasing response with complexity during habituation, and decreasing response with complexity to deviants.

All these areas have been shown to be associated with temporal sequence processing, although mostly with oddball paradigms using much shorter or simpler sequences (Bekinschtein et al., 2009; Huettel et al., 2002; Planton & Dehaene, 2021; Wang et al., 2015, 2019). They can be decomposed into modality-specific and modality-independent regions (Frost et al., 2015). STG activation was observed for auditory sequences here and in other work (Bekinschtein et al., 2009; Wang et al., 2015) but not visuo-spatial ones (Wang et al., 2019). The modality specificity of STG was explicitly confirmed by Planton and Dehaene (2021) using visual and auditory sequences with identical structures. Other regions, meanwhile, were modality-independent and coincided with those found in a similar paradigm with visuo-spatial sequences (Wang et al., 2019), consistent with a role in abstract rule formation. The IPS and preCG, in particular, are jointly activated in various conditions of mental calculation and mathematics (Amalric & Dehaene, 2017; Dehaene et al., 2003), with anterior IPS housing a representation of number (Dehaene et al., 2003; Eger et al., 2009; Harvey et al., 2013; Kanayet et al., 2018). The overlap between auditory sequences and arithmetic was confirmed here using sensitive single-subject analyses (Figure 5). PreCG and IPS may thus be jointly involved in the nested “for i=1:n” loops of the proposed language, and in the real-time tracking of item and chunk number needed to follow a given auditory sequence even after it was learned. While here they coactivated with STG, in proportion to LoT complexity, in a visuo-spatial version of the present task they did so together with bilateral occipito-parietal areas (Wang et al., 2019). This is consistent with the behavioral observation that the very same language, involving concatenation, loops and recursion, when applied to distinct visual or auditory primitives, can account for both domains (Dehaene et al., 2022a; Planton et al., 2021).

Our data also point to the SMA, or rather pre-SMA (Nachev et al., 2008), in processing increasingly complex sequences. Such a domain-general sequence processing function was indeed advocated by Cona & Semenza (2017) given its various involvements in action sequences, music processing, numerical cognition, spatial processing, time processing, as well as language. Remarkably, cerebellum also participates in our complexity network. Its role in working memory has been rarely reported or discussed and might have been underestimated in the parsing of non-motor sequences, as it is classically associated with motor sequence learning (Jenkins et al., 1994; Toni et al., 1998). The present results confirms that the cerebellum may be involved in abstract, non-motor sequence encoding and expectation (Leggio et al., 2008; Molinari et al., 2008; Nixon, 2003). Indeed, cerebellum, SMA and premotor cortex were already reported as involved in the passive listening of rhythms (J. L. Chen et al., 2008), consistent with a role in the identification of sequence regularities. A tentative hypothesis is that (pre)SMA, cerebellum and possibly premotor cortex may participate in a beat- (Morillon & Baillet, 2017) or time-processing network (Coull et al., 2011), thus possibly involved in the translation from the abstract structures of the proposed language to concrete, precisely timed sensory predictions.

Interestingly, we found that, while task performance was strictly linearly related to LoT complexity, fMRI activity did not. Rather, as the sequence becomes too complex, activation tended to stop increasing, or even decreased, just yielding a significant downward quadratic trend. Wang et al. (2019) observed a similar effect with visuo-spatial sequences. In both cases, we ensured that the highest complexity sequences did not have any significant regularity in our language and, given their length, couldn't be easily memorized. The collapse of activity for maximum LoT complexity, in regions that are precisely involved in working memory is therefore logical. Indeed, in a more classical object memory task, Vogel and Machizawa (2004) found that working memory activity does not solely increase with the number of elements stored in working memory, but saturates or decreases when the storage capacity, thought to be around three or four items (Cowan, 2001) is exceeded. Naturally, such a collapse can only lead to reduced predictions and therefore reduced violation detection – thus explaining that fMRI, MEG and behavioral responses to deviants vary linearly with complexity, while model-related fMRI activations vary as an inverted U function of complexity. An analogous phenomenon was described in infants (Kidd et al., 2012, 2014): they allocate their attention to visually or auditory presented sequences that are neither too simple nor too complex, thus showing a U-shaped pattern that implies boredom for stimuli with low information content and saturation from stimuli that exceed their cognitive resources.

Detailed examination of the responses to violations in MEG confirmed that human participants were able to encode details of the hierarchical structures of sequences. Not only did the amplitude of violation responses tightly track the proposed LoT complexity (Figure 7), but the specific violation responses proved that the human brain changed its expectations in a hierarchical manner (Figure 8). This was clearest in the case of the Pairs+Alt1 sequence, which consists in 2 pairs (AABB) followed by 4 alternations (ABAB). In those two consecutive parts, the predictions are exactly opposite at central locations (A**AB**B versus A**BA**B), such that what is a violation for one is a correct prediction for the other, and vice-versa. The fact that we observe significant violation responses at each of these locations (i.e. locations 10, 12, 14 and 15 in the pairs+alt1 sequence), as well as for the matched *Alternate* and *Pairs* sequences, indicates that the human brain is able to quickly change its anticipations as a function of sequence hierarchical structure. To do so, it must contain a representation of sequences as nested parts with parts, and switch between those parts after a fixed number of items (4 in this case). Violation detection in the *Pairs* and *Quadruplets* sequences further confirmed that subjects kept track of the exact number of items in each subsequence, since their brain reacted to violations which either shortened or, on the contrary, lengthened a chunk of identical consecutive items.

While present and past results thus indicate that a language is necessary to account for the human encoding of binary auditory sequences (Dehaene et al., 2022a; Planton et al., 2021), this language appears to differ from those used for communication, since it involves repetitions, numbers and symmetries, while the syntax of natural language systematically avoids these features (Moro, 1997; Musso et al., 2003). In agreement with this observation, there was little overlap between our auditory sequence complexity network and the classical left-hemisphere language network. Instead, complexity effects were systematically distributed symmetrically in both hemispheres, unlike natural language processing. Within individually-defined language fROIs (defined by their activity during visual or auditory sentence processing relative to a low-level control), no significant complexity effect was found except in a single region, the left IFGoper (a negative effect of complexity for deviants was also found there and in pSTS). Even that finding may well be a partial volume effect, as this area was absent from whole-brain contrasts, and the centroid of the complexity-related activation was centered at a more dorsal location in preCG (Figure 3). Broca's area is the main candidate region for language-like processing of hierarchical structures, and such role is advocated for in various previous rule-learning studies using artificial grammars (Bahlmann et al., 2008; Fitch & Friederici, 2012; Friederici et al., 2006) structured sequences of actions (Badre & D'Esposito, 2007; Koechlin & Jubault, 2006), sequence processing (Wang et al., 2015), and even music (Maess et al., 2001; Patel, 2003). However, Broca's area is a heterogeneous region (Amunts et al., 2010), of which certain sub-regions support language while others underlie a variety of other cognitive functions, including mathematics and working memory (Fedorenko et al., 2012). Interpretation must remain careful since functions that were once thought to overlap in Broca's area, such as language and musical syntax (Fadiga et al., 2009; Koelsch et al., 2002; Kunert et al., 2015), are now clearly dissociated by higher-resolution single-subject analyses (X. Chen et al., 2021).

Conversely, a very different picture was observed when examining the overlap of LoT complexity fMRI activity and the mathematical calculation network. There was considerable overlap at the whole-brain level (SMA, IPS, premotor cortex, cerebellum) and, most importantly, a significant sequence complexity effect within each of the individual mathematical fROIs. A similar result was reported by Wang et al. (2019); they found activation of mathematics-related regions but not language-related ones when participants were processing visuo-spatial sequences. Planton and Dehaene (2021) actually already reached a similar conclusion by showing novelty effects to pattern violations of both visual and auditory short sequences in mathematics but not in language areas. Since theirs, as well as the present data, was obtained with binary sequences which, contrary to Wang et al. (2019) were devoid of geometrical content, overall those results that the amodal language of thought that encodes sequence pattern shares common neural mechanisms with mathematical thinking.

The present results therefore support the hypothesis that the human brain hosts multiple internal languages, depending on the types of structures and contents that are being processed (Dehaene et al., 2022a; Fedorenko & Varley, 2016; Hagoort, 2013). While the capacity to encode nested sequences may well be a fundamental overarching function of the human brain, fundamental to the manipulation of hierarchical structures in language, mathematics, music, complex actions, etc. (Dehaene et al., 2015; Fitch, 2014; Hauser et al., 2002; Lashley, 1951), those abilities may rely on partially dissociable networks. This conclusion fits with much prior evidence that, at the individual level, language and mathematics do not share the same cerebral substrates and may be dissociated by brain injuries (Amalric & Dehaene, 2016, 2017; Fedorenko & Varley, 2016), just like language and music (J. L. Chen et al., 2008; Norman-Haignere et al., 2015; Peretz et al., 2015). During hominization, an enhanced functionality for recursive nesting may have jointly emerged in all of those neuronal circuits. In the future, this hypothesis could be tested by submitting non-human primates to the present hierarchy of sequences, and examine up to which level their brains can react to violation. We already know that the macaque monkey brain can detect violations of simple habitual, sequential or numerical patterns (Uhrig et al., 2014; Wilson et al., 2013), with both convergence (Wilson et al., 2017) and divergence (Wang et al., 2015) relative to human results. The present design may help determine precisely where to draw the line.

# Materials and methods

## Participants

Nineteen participants (10 men, $M_{age}$ = 27.6 years, $SD_{age}$ = 4.7 years) took part in the MEG experiment and twenty-three (11 men, $M_{age}$ = 26.1 years, $SD_{age}$ = 4.7 years) in the fMRI experiment. We did not test any effect of gender on the results of this study. All participants had normal or corrected to normal vision and no history or indications of psychological or neurological disorders. In compliance with institutional guidelines, all subjects gave written informed consent prior to enrollment and received 90€ as compensation. The experiments were approved by the national ethical committees (CPP Ile-de-France III and CPP Sud-Est VI).

## Stimuli and tasks

Auditory binary sequences of 16 sounds were used in both experiments. They were composed of low pitch and high pitch sounds, constructed as the superimposition of sinusoidal signals of respectively f = 350Hz, 700Hz and 1400Hz, and f = 500Hz, 1000Hz and 2000Hz. Each tone lasted 50 ms and the 16 tones were presented in sequence with a fixed SOA of 250ms.

Ten 16-items sequential patterns spanning a large range of complexities were selected (see Figure 1A). Six of them were used in previous behavioral experiments (Planton et al., 2021). The complexity metric used to predict behavior and brain activity was the "Language-of-thought – *chunk*" complexity, which was previously shown to be well correlated with behavior (Planton et al., 2021). This metric roughly measures the length of the shortest description of the pattern in a formal language that uses a small set of atomic rules (e.g. repetition, alternation) that can be recursively embedded. The *chunk* version of the metric includes only expressions that preserve chunks of consecutive repeated items (for instance, the sequence ABBA is parsed as [A][BB][A] rather than [AB][BA]). 10 sequences were used in the fMRI experiment, and 7 of them in the MEG experiment (i.e. all but *Pairs&Alt.2*, *ThreeTwo* and *CenterMirror*).

Each auditory sequence (4000 ms long) was repeatedly presented to a participant in a mini-session with 500 ms ITI. Mini-sessions had the following structure. Participants first discovered and encoded the sequence during a habituation phase of 10 trials. Then, during a test phase, occasional violations consisting in the replacement of a high pitch sound by a low pitch one (or vice-versa) were presented at the locations specified in Figure 1A. As described in Figure 1B, in the MEG experiment, the test phase included 36 trials of which 2/3 comprised a deviant sound. In the fMRI experiment, the test phase included 18 trials of which 1/3 comprised a deviant sound. Participants were unaware of the mini-session structure.

In the MEG experiment, habituation and test sequences followed each other seamlessly, and participants were merely asked to listen attentively. After each mini-session, they were asked one general question about what they had just heard such as: *How many different sounds could you hear? Did you find it musical? How complex was the sequence of sounds?* The full experiment was divided temporally into 2 parts such that the 7 sequence types appeared twice, once in each version (starting with A or B), once at the beginning and once at the end of the experiment. The overall experiment lasted about 80 minutes.

In the fMRI experiment, participants were explicitly instructed to detect and respond to violations, by pressing a button, as quickly as possible, with either their right or left hand. The correct response button (left or right, counterbalanced over the two repetitions of each sequence) was indicated by a 2s visual message on the screen during the rest period preceding the first test trial. In order to optimize the estimation of the BOLD response, trials were presented in two blocks of 5 trials for the habituation phase, then three blocks of 6 trials for the test phase, separated by rest periods of variable duration (6s ± 1.5). The 10 sequences appeared twice, once in each version (starting with A or B). The 20 mini-sessions were presented across 5 fMRI sessions of approximately 11 minutes.

## Post-experimental sequence bracketing task

After the experiment, participants were given a questionnaire to assess their own representation of the structure of the sequence. For each sequence of the experiment (i.e. 7 for the MEG participants, 10 for the fMRI participants), after listening to it several times if needed, participants were asked to segment the sequence by drawing brackets (opening and closing) on its visual representation As and Bs were respectively represented by empty and filled circles on a sheet of paper). In this way, they were instructed to indicate how they tended to group consecutive items together in their mind when listening to the sequence, if they did.

## MEG experiment procedures

### MEG recordings

Participants listened to the sequences 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 0.1Hz. The data was then resampled at 250 Hz.

Eye movements and heartbeats were monitored with vertical and horizontal electro-oculograms (EOGs) and electrocardiograms (ECGs). Head shape was digitized using various points on the scalp as well as the nasion, left and right pre-auricular points (FASTTRACK, Polhemus). 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. Sounds were presented using Eatymotic audio system (an HiFi-quality artifact-free headphone system with wide frequency response) while participants had to fixate a central cross. The analysis was performed with MNE Python (Gramfort et al., 2013; Jas et al., 2018), version 0.23.0.

#### *Data cleaning: Maxfiltering*

We applied the signal space separation algorithm *mne.preprocessing.maxwell_filter* (Taulu et al., 2004) to suppress magnetic signals from outside the sensor helmet and interpolate bad channels that we identified visually in the raw signal and in the power spectrum. This algorithm also compensated for head movements between experimental blocks by realigning all data to an average head position.

#### *Data cleaning: ICA*

Oculomotor and cardiac artefacts were removed performing an independent component analysis (ICA) on the four last runs of the experiment. 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. A maximum of 1 component for the cardiac artefact and 2 components for the ocular artefacts were considered. Finally, we removed them from the whole recording (14 runs).

### *Data cleaning: Autoreject*

We used an automated algorithm for rejection and repair of bad trials (Jas et al., 2017) that computes the optimal peak-to-peak threshold per channel-type in a cross-validated manner. It was applied to baselined epochs and removed on average 4.6% of the epochs.

### *Epoching parameters and projection on magnetometers*

Epochs on items were baselined from -50 ms to 0 ms (stimulus onset) and epochs on the full sequences were baselined between -200ms to 0ms (first sequence item onset). For sensor level analyses, instead of working with the 306 sensors (102 magnetometers and 206 gradiometers), we projected the spherical sources of signal onto the magnetometers using MNE epochs method *epochs.as_type('mag',mode='accurate').*

## Univariate analyses

### *GFP and linear regressions*

Global field power was computed as the root-mean-square of evoked responses or the difference of evoked responses. Linear regressions were computed using 4-fold cross-validation and with the *linear_model.LinearRegression* function of scikit-learn package version 0.24.1. Pearson correlation was computed with the *stats.pearsonr* function from *scipy* package. The predictors for surprise from transition probabilities were computed using an ideal observer Bayesian model learning first-order transitions with an exponential memory decay over 100 items. This was done thanks to the TransitionProbModel python package, which is the python version of the *Matlab version* used in (Maheu et al., 2019; Meyniel et al., 2016).

### *Source reconstruction*

A T1-weighted anatomical MRI image with 1 mm isometric resolution was acquired for each participant (3T Prisma Siemens scanner). The anatomical MRI was segmented with FreeSurfer (Dale et al., 1999; Fischl et al., 2002) and co-registered with MEG data in MNE using the digitized markers. A three-layer boundary element model (inner skull, outer skull and outer skin) was used to estimate the current-source density distribution over the cortical surface. Source reconstruction was performed on the linear regression coefficients  using the dSPM solution with MNE default values (loose orientation of 0.2, depth weighting of 0.8, SNR value of 3) (Dale et al., 2000). The noise covariance matrix used for data whitening was estimated from the signal within the 200 ms preceding the onset of the first item of each sound sequence. The resulting sources estimates were transformed to a standard anatomical template (*fsaverage*) with 20484 vertices using the MNE morphing procedure, and averaged across subjects.

## Multivariate analyses

Data was smoothed with a 100 ms sliding window and, instead of working with the 306 sensors (102 magnetometers and 206 gradiometers), we projected the spherical sources of signal onto the magnetometers using MNE epochs method *epochs.as_type('mag',mode='accurate').*

### *Time-resolved multivariate decoding of brain responses to standard and deviant sounds*

The goal of multivariate of time-resolved decoding analyses was to predict from single-trial brain activity (*X*) a specific categorical variable (*y*), namely if the trial corresponded to the presentation of a deviant sound or not. These analyses were performed following King et al's preprocessing pipeline (King & Dehaene, 2014) using MNE-python (Gramfort et al., 2013). Prior to model fitting, channels were z-scored across trials for every time-point. The 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).

### *Cross-validation*

One run was dedicated to each version of the sequence (7 sequence types x 2 versions [starting with A or starting with B] = 14 runs). To build the training set, we randomly picked one run for each sequence, irrespectively of the sequence version. We trained the decoder on all deviant trials of the 7 sequences and on standard trials (non-deviant trials from the test phase) that were matched to sequence-specific deviants in ordinal position. We then tested this decoder on the remaining 7 blocks, determining its performance for the 7 sequences separately. The training and the testing sets were then inverted, resulting in a 2-folds cross-validation. This procedure avoided any confound with item identity, as the sounds A and B were swapped in the cross-validation folds.

#### *Generalization across time*

To access the temporal organization of the neural representations, we computed the generalization-across-time (GAT) matrices (King & Dehaene, 2014). 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).

#### *Statistical analyses*

Temporal, spatiotemporal and temporal-temporal cluster-based permutation tests were computed on the time-windows of interest (0-350ms for habituation and standard items and 0-600ms for deviants) using *stats.permutation_cluster_1samp_test* from MNE python package. To compute spatiotemporal clusters, we provided the function with an adjacency matrix from *mne.channels.find_ch_connectivity.*

## fMRI experiment procedures

### Localizer session

Together with the main sequence processing task described above, the fMRI experimental protocol also included a 6-min localizer session, designed to localize cerebral regions involved in language processing and in mathematics. It was derived from a previously published functional localizer (see Pinel et al., 2007, for details) and was already used elsewhere (Planton & Dehaene, 2021). A sentence processing network was identified in each subject by contrasting sentence reading/listening conditions (i.e. visually and auditorily presented sentences) from control conditions (i.e. meaningless auditory stimuli consisting in rotated sentences, and meaningless visual stimuli of the same size and visual complexity as visual words). A mathematics network was identified in each subject by contrasting mental calculation conditions (i.e. mental processing of simple subtraction problems, such as 7 − 2, presented visually, and auditorily) from sentence reading/listening conditions.

### fMRI acquisition and preprocessing

MRI acquisition was performed on a 3T scanner (Siemens, Tim Trio), equipped with a 64-channel head coil. 354 functional scans covering the whole brain were acquired for each of the 5 sessions of the main experiment, as well as 175 functional scans for the localizer session, all using a T2*-weighted gradient echo-planar imaging (EPI) sequence (69 interleaved slices, TR = 1.81 s, TE = 30.4 ms, voxel size = 1.75 mm3, multiband factor = 3). To estimate distortions, two volumes with opposite phase encoding direction were acquired: one volume in the anterior to posterior direction (AP) and one volume in the other direction (PA). A 3D T1-weighted structural image was also acquired (TR = 2.30 s, TE = 2.98 ms, voxel size = 1.0 mm3).

Data processing (except the TOPUP correction) was performed with SPM12 (Wellcome Department of Cognitive Neurology, http://www.fil.ion.ucl.ac.uk/spm). The anatomical scan was spatially normalized to a standard Montreal Neurological Institute (MNI) reference anatomical template brain using the default parameters. Functional images were unwarped (using the AP/PA volumes, processed with the TOPUP software; FSL, fMRIB), corrected for slice timing differences (first slice as reference), realigned (registered to the mean using 2nd degree B-Splines), coregistered to the anatomy (using Normalized Mutual Information), spatially normalized to the MNI brain space (using the parameters obtained from the normalization of the anatomy), and smoothed with an isotropic Gaussian filter of 5-mm FWHM.

In addition to the 6 motion regressors from the realignment step, 12 regressors were computed using the aCompCor method (Behzadi et al., 2007), applied to the CSF and to white matter (first 5 components of two principal component analyses, and 1 for the raw signal), in order to better correct for motion-related and physiological noise in the statistical models (using the PhysIO Toolbox, Kasper et al., 2017). Additional regressors for motion outliers were also computed (framewise displacement larger than 0.5 mm; see Power et al., 2012), they represented 0.5% of volumes per subject on average. One participant was excluded from the fMRI analyses due to excessive movement in the scanner (average translational displacement of 2.9 mm within each fMRI session, which was 3.3 SD above group average).

## fMRI analysis

### *General linear model*

Statistical analyses were performed using SPM12 and general linear models (GLM) that included the motion-related and physiological noise-related regressors (described above) as covariates of no interest. fMRI images were high-pass filtered at 0.01 Hz. Time series from the sequences of stimuli of each condition (each tone modeled as an event) were convolved with the canonical hemodynamic response function (HRF). Specifically, for each of the twenty mini-sessions (i.e. each sequence being tested twice, reverting the attribution of the two tones), one regressor for the items of the habituation phase, one for the items of the test phase and one for the deviant items were included in the GLM. Since motor responses and deviant trials were highly collinear, manual motor responses were not modeled. However, motor responses could be less frequent for more complex sequences (i.e. increased miss rate), thus creating a potential confound with the effect of complexity in deviant trials. We thus also computed an alternative model in which only correctly-detected deviants trials were included. In order to test for a relationship between brain activation and LoT complexity in different trial types (i.e. habituation trials, deviant trials), corresponding beta maps for each of the 10 sequences and each participant were entered in second-level within-subject ANOVA analyses. Linear parametric contrasts using the LoT complexity value were then computed.

### *Cross-validated ROI analyses*

To further test the reliability of the complexity effect across participants, a cross-validated region-of-interest (ROI) analysis, using individually-defined functional ROIs (fROIs), was conducted. Nine of the most salient peaks from the positive LoT complexity contrast in habituation were first selected, and used to build nine 20-mm-diameter spherical search volumes: supplementary motor area (SMA; coordinates: -1, 5, 65), right precentral gyrus (R-preCG; 46, 2, 44), left precentral gyrus (L-preCG; -47, 0, 45), right intraparietal sulcus (R-IPS; 36, -46, 56), left intraparietal sulcus (L-IPS; -31, -42, 44), right superior temporal gyrus (R-STG; 48, -32, 3), left superior temporal gyrus (L-STG; -68, -23, 5), lobule VI of the left cerebellar hemisphere (L-CER6; -29, -56, -28) and lobule VI of the right cerebellar hemisphere (R-CER8; 22, -68, -51). Individual fROIs were then defined for each participant by selecting the 20% most active voxels at the intersection between each search volume and the contrast "LoT complexity effect in habituation" computed on half of the blocks (i.e. blocks of sequences starting with "A"). Mean contrast estimates for each fROI and each condition was then extracted using the other half of the blocks (i.e. blocks of sequences starting with "B"). The same procedure was repeated a second time by reversing the role of the two halves (i.e. fROIs computed using blocks of sequences starting with "B", data extracted from blocks of sequences starting with "A"). To test for the significance of the complexity effect in each ROI, the mean of the output of the two procedures (i.e. the cross-validated activation value), for each of the ten conditions (i.e. habituation blocks for each of the 10 sequences) and each participant, was entered in a linear mixed model mixed effect model with participant as random factor and LoT complexity value as a fixed effect predictor. P values were corrected for multiple comparison using Bonferroni correction for nine ROIs. Along with such linear effect of complexity, we also tested a quadratic effect, by adding a quadratic term in the mixed effect model.

In order to track activation over time, we also extracted, using the same cross validated procedure, the BOLD activation time course for each 28-trials mini-session. To account for the fact that the duration of rest periods between blocks could vary, data were actually extracted for a [-6s – 32s] period relative to the onset of the first trial of each block rest period, and the whole mini-session curve was recomposed by averaging over the overlapping period of two consecutive parts (see vertical shadings in Figure 4A). Each individual time course was upsampled and smoothed using cubic spline interpolation, and baseline-corrected with a 6-seconds period preceding the onset of the first trial.

Finally, two set of ROIs were selected in order to test for the involvement of language and mathematics-related areas in the present sequence processing task, and especially to assess a potential sequence complexity effect. Seven language-related ROIs came from the sentence processing experiment of Pallier et al. (2011): pars orbitalis (IFGorb), triangularis (IFGtri), and opercularis (IFGoper) of the inferior frontal gyrus, temporal pole (TP), temporoparietal junction (TPJ), anterior superior temporal sulcus (aSTS) and posterior superior temporal sulcus (pSTS). Seven mathematics-related ROIs came from the mathematical thinking experiment of Amalric & Dehaene (2016): left and right intraparietal sulcus (IPS), left and right superior frontal gyrus (SFG), left and right precentral/inferior frontal gyrus (preCG/IFG), supplementary motor area (SMA). These two set of ROIs were already used in the past (Planton & Dehaene, 2021; Wang et al., 2019). In order to build individual and functional ROIs from these literature-based ROIs, we used the same procedure as Planton & Dehaene (2021) consisting in selecting, for each subject, the 20% most active voxels within the intersection between the ROI mask and an fMRI contrast of interest from the independent localizer session. The contrast of interest was "Listening & reading sentences > Rotated speech & false font script" for the ROIs of the language network, and "Mental calculation visual & auditory > Sentence listening & reading" for the ROIs of the mathematics network. Mean contrast estimates for each fROI and each condition was then extracted, and entered into linear mixed model mixed effect model with participant as random factor and LoT complexity value as a fixed effect predictor. A Bonferonni correction for 14 ROIs was applied to the p values.

## Behavioral data analysis

Data for the sequence bracketing task included all productions collected in the fMRI and MEG experiment (42 participants for seven sequences, and 23 participants for the three that were only presented during fMRI). For each production, we counted the total number of brackets (opening and closing) drawn at each interval between two consecutive items (as well as before the first and after the last item, resulting in a vector of length 17) (see Figure 2A). To determine if participants' reported sequence structure matched the predictions of the LoT model, we computed the correlation between the average over participants of the number of brackets in each interval and the postulated bracketing of the sequence (derived from its expression in the LoT). For the first two sequences, the representations "[A][A][A]…" and "[AAA…]", as well as "[A][B][A]…" and "[ABA…]", respectively derived from the expressions [+0]^16 and [+0]^16<b>, were considered as equivalent.

For the violation detection task of the fMRI experiment, we considered as a correct response (or "hit") all button presses occurring between 200 ms and 2500 ms after the onset of a deviant sound. We thus allowed for potential delayed responses (but found that 97.7% of correct responses were below 1500 ms). An absence of response in this interval was counted as a miss, a button press outside this interval was counted as a false alarm. We then computed, for each subject and each sequence, the average response time as well as, using the proportions of hits and false alarms, the sensitivity (or d'). The method of Hautus (1995) was used to adjust extreme values. In order to test whether subject performance was predicted by LoT complexity, we performed linear regressions on group-averaged data, as well linear mixed models including participant as random factor on the by-subject data. Analyses were performed in R 4.0.2 (R Core Team, 2020), using the lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017) packages. Surprise for each deviant item was computed from transition probabilities, within each block for each subject, using an ideal observer Bayesian model (Maheu et al., 2019; Meyniel et al., 2016), and tested as an additional predictor in the mixed effect models. For the analysis of d', we used the average surprise of the deviant items of the block (i.e. all deviants presented to the subject, whether or not they detected them). For the analysis of response times, we used the average surprise of the correctly-detected deviant items of the block

# References

Aksentijevic, A., & Gibson, K. (2012). Complexity equals change. *Cognitive Systems Research*, *15–16*, 1–16. https://doi.org/10.1016/j.cogsys.2011.01.002

Al Roumi, F., Marti, S., Wang, L., Amalric, M., & Dehaene, S. (2021b). Mental compression of spatial sequences in human working memory using numerical and geometrical primitives. *Neuron*, *109*(16), 2627-2639.e4. https://doi.org/10.1016/j.neuron.2021.06.009

Alexander, C., & Carey, S. (1968). Subsymmetries. *Perception & Psychophysics*, *4*(2), 73–77. https://doi.org/10.3758/BF03209511

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proceedings of the National Academy of Sciences*, *113*(18), 4909–4917. https://doi.org/10.1073/pnas.1603205113

Amalric, M., & Dehaene, S. (2017). Cortical circuits for mathematical knowledge: Evidence for a major subdivision within the brain's semantic networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20160515–20160515. https://doi.org/10.1098/rstb.2016.0515

Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., & Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers. *PLOS Computational Biology*, *13*(1), e1005273. https://doi.org/10.1371/journal.pcbi.1005273

Amunts, K., Lenzen, M., Friederici, A. D., Schleicher, A., Morosan, P., Palomero-Gallagher, N., & Zilles, K. (2010). Broca's region: Novel organizational principles and multiple receptor mapping. *PLoS Biol*, *8*(9). https://doi.org/10.1371/journal.pbio.1000489

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), Article 10. https://doi.org/10.1038/nrn1201

Baddeley, A. D., & Hitch, G. (1974). Recent advances in learning and motivation. In *Working Memory, Vol. 8,*.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2008.02.004

Badre, D., & D'Esposito, M. (2007). Functional Magnetic Resonance Imaging Evidence for a Hierarchical Organization of the Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *19*(12), 2082–2099. https://doi.org/10.1162/jocn.2007.19.12.2082

Badre, D., Kayser, A. S., & D'Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron*, *66*(2), 315–326. https://doi.org/10.1016/j.neuron.2010.03.025

Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, *42*(2), 525–534. https://doi.org/10.1016/j.neuroimage.2008.04.249

Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proceedings of the National Academy of Sciences*, *113*(5), E616–E625. https://doi.org/10.1073/pnas.1508523113

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677. https://doi.org/10.1073/pnas.0809667106

Bendixen, A., Roeber, U., & Schröger, E. (2007). Regularity extraction and application in dynamic auditory stimulus sequences. *Journal of Cognitive Neuroscience*, *19*(10), 1664–1677. https://doi.org/10.1162/jocn.2007.19.10.1664

Bendixen, A., Schröger, E., & Winkler, I. (2009). I Heard That Coming: Event-Related Potential Evidence for Stimulus-Driven Prediction in the Auditory System. *The Journal of Neuroscience*, *29*(26), 8447–8451. https://doi.org/10.1523/JNEUROSCI.1493-09.2009

Bhanji, J. P., Beer, J. S., & Bunge, S. A. (2010). Taking a gamble or playing by the rules: Dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. *NeuroImage*, *49*(2), 1810–1819. https://doi.org/10.1016/j.neuroimage.2009.09.030

Botvinick, M., & Watanabe, T. (2007). From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *Journal of Neuroscience*, *27*(32), 8636–8642. https://doi.org/10.1523/JNEUROSCI.2110-07.2007

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage*, *44*(2), 509–519. https://doi.org/10.1016/j.neuroimage.2008.09.015

Chao, Z. C., Takaura, K., Wang, L., Fujii, N., & Dehaene, S. (2018). Large-Scale Cortical Networks for Hierarchical Prediction and Prediction Error in the Primate Brain. *Neuron*, *100*(5), 1252-1266.e3. https://doi.org/10.1016/j.neuron.2018.10.004

Chater, N., & Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science? *Trends in Cognitive Sciences*, *7*(1), 19–22. https://doi.org/10.1016/S1364-6613(02)00005-0

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex (New York, N.Y.: 1991)*, *18*(12), 2844–2854. https://doi.org/10.1093/cercor/bhn042

Chen, X., Affourtit, J., Ryskin, R., Regev, T. I., Norman-Haignere, S., Jouravlev, O., Malik-Moraleda, S., Kean, H., Varley, R., & Fedorenko, E. (2021). *The human language system does not support music processing* (p. 2021.06.01.446439). https://doi.org/10.1101/2021.06.01.446439

Cona, G., & Semenza, C. (2017). Supplementary motor area as key structure for domain-general sequence processing: A unified account. *Neuroscience and Biobehavioral Reviews*, *72*, 28–42. https://doi.org/10.1016/j.neubiorev.2016.10.033

Coull, J. T., Cheng, R.-K., & Meck, W. H. (2011). Neuroanatomical and Neurochemical Substrates of Timing. *Neuropsychopharmacology*, *36*(1), Article 1. https://doi.org/10.1038/npp.2010.113

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2010). The Magical Mystery Four: How Is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57. https://doi.org/10.1177/0963721409359277

Dehaene, S., Al Roumi, F., Lakretz, Y., Planton, S., & Sablé-Meyer, M. (2022b). Symbols and mental programs: A hypothesis about human singularity. *Trends in Cognitive Sciences*, *26*(9), 751–766. https://doi.org/10.1016/j.tics.2022.06.010

Dehaene, S., Meyniel, F., Wacongne, C., Wang, L., & Pallier, C. (2015). The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron*, *88*(1), 2–19. https://doi.org/10.1016/j.neuron.2015.09.019

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Delahaye, J.-P., & Zenil, H. (2012). Numerical evaluation of algorithmic complexity for short strings: A glance into the innermost structure of randomness. *Applied Mathematics and Computation*, *219*(1), 63–77. https://doi.org/10.1016/j.amc.2011.10.006

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009). Deciphering cortical number coding from human brain activity patterns. *Curr Biol*, *19*(19), 1608–1615. https://doi.org/10.1016/j.cub.2009.08.047

El Karoui, I., King, J.-R., Sitt, J., Meyniel, F., Van Gaal, S., Hasboun, D., Adam, C., Navarro, V., Baulac, M., Dehaene, S., Cohen, L., & Naccache, L. (2015). Event-Related Potential, Time-frequency, and Functional Connectivity Facets of Local and Global Auditory Novelty Processing: An Intracranial Study in Humans. *Cerebral Cortex*, *25*(11), 4203–4212. https://doi.org/10.1093/cercor/bhu143

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Ann N Y Acad Sci*, *1169*, 448–458. https://doi.org/10.1111/j.1749-6632.2009.04582.x

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Current Biology: CB*, *22*(21), 2059–2062. https://doi.org/10.1016/j.cub.2012.09.011

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: Evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132–153. https://doi.org/10.1111/nyas.13046

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. *Nature*, *407*(6804), 630–633. https://doi.org/10.1038/35036586

Ferrigno, S., Cheyette, S. J., Piantadosi, S. T., & Cantlon, J. F. (2020). Recursive sequence generation in monkeys, children, U.S. adults, and native Amazonians. *Science Advances*, *6*(26), eaaz1002. https://doi.org/10.1126/sciadv.aaz1002

Fitch, W. T. (2004). Computational Constraints on Syntactic Processing in a Nonhuman Primate. *Science*, *303*(5656), 377–380. https://doi.org/10.1126/science.1089401

Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, *11*(3), 329–364. https://doi.org/10.1016/j.plrev.2014.04.005

Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: An overview. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1933–1955. https://doi.org/10.1098/rstb.2012.0103

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*, 87–104. https://doi.org/10.1111/nyas.12406

Fló, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J., & Ferry, A. L. (2019). Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Developmental Science*, *22*(4), e12802. https://doi.org/10.1111/desc.12802

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard university press.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(7), 2458–2463. https://doi.org/10.1073/pnas.0509389103

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Fujii, N., & Graybiel, A. M. (2003). Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science (New York, N.Y.)*, *301*(5637), 1246–1249. https://doi.org/10.1126/science.1086872

Gauvrit, N., Zenil, H., Delahaye, J.-P., & Soler-Toscano, F. (2014). Algorithmic complexity for short binary strings applied to psychology: A primer. *Behavior Research Methods*, *46*(3), 732–744. https://doi.org/10.3758/s13428-013-0416-0

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, *440*(7088), 1204–1207. https://doi.org/10.1038/nature04675

Glanzer, M., & Clark, W. H. (1963). Accuracy of perceptual recall: An analysis of organization. *Journal of Verbal Learning and Verbal Behavior*, *1*(4), 289–299. https://doi.org/10.1016/S0022-5371(63)80008-0

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00267

Grunwald, P. (2004). A tutorial introduction to the minimum description length principle. *ArXiv:Math/0406077*. http://arxiv.org/abs/math/0406077

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00416

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic Representation of Numerosity in the Human Parietal Cortex. *Science*, *341*(6150), 1123–1126. https://doi.org/10.1126/science.1239052

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, *298*(5598), 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*(3), B53-64. https://doi.org/10.1016/s0010-0277(00)00132-3

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values ofd′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? *Neuroscience*, *389*, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nat Neurosci*, *5*(5), 485–490.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2), 339. https://doi.org/10.1037/a0034221

Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage*, *159*, 417–429. https://doi.org/10.1016/j.neuroimage.2017.06.030

Jas, M., Larson, E., Engemann, D. A., Leppäkangas, J., Taulu, S., Hämäläinen, M., & Gramfort, A. (2018). A Reproducible MEG/EEG Group Study With the MNE Software: Recommendations, Quality Assessments, and Good Practices. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00530

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *14*(6), 3775–3790.

Jiang, X., Long, T., Cao, W., Li, J., Dehaene, S., & Wang, L. (2018). Production of Supra-regular Spatial Sequences by Macaque Monkeys. *Current Biology: CB*, *28*(12), 1851-1859.e4. https://doi.org/10.1016/j.cub.2018.04.047

Kanayet, F. J., Mattarella-Micke, A., Kohler, P. J., Norcia, A. M., McCandliss, B. D., & McClelland, J. L. (2018). Distinct Representations of Magnitude and Spatial Position within Parietal Cortex during Number–Space Mapping. *Journal of Cognitive Neuroscience*, *30*(2), 200–218. https://doi.org/10.1162/jocn_a_01199

Karuza, E. A., Kahn, A. E., & Bassett, D. S. (2019). Human Sensitivity to Community Structure Is Robust to Topological Variation. *Complexity*, *2019*, e8379321. https://doi.org/10.1155/2019/8379321

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS One*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks Effect in Infant Auditory Attention. *Child Development*, *85*(5), 1795–1804. https://doi.org/10.1111/cdev.12263

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: The temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/J.TICS.2014.01.002

King, J. R., Faugeras, F., Gramfort, A., Schurger, A., El Karoui, I., Sitt, J. D., Rohaut, B., Wacongne, C., Labyt, E., Bekinschtein, T., Cohen, L., Naccache, L., & Dehaene, S. (2013). Single-trial decoding of auditory novelty responses facilitates the detection of residual consciousness. *NeuroImage*, *83*, 726–738. https://doi.org/10.1016/j.neuroimage.2013.07.013

Kóbor, A., Takács, Á., Kardos, Z., Janacsek, K., Horváth, K., Csépe, V., & Nemeth, D. (2018). ERPs differentiate the sensitivity to statistical probabilities and the learning of sequential structures during procedural learning. *Biological Psychology*, *135*, 180–193. https://doi.org/10.1016/j.biopsycho.2018.04.001

Koechlin, E., & Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, *50*(6), 963–974. https://doi.org/10.1016/j.neuron.2006.05.017

Koechlin, E., Ody, C., & Kouneiher, F. (2003). The Architecture of Cognitive Control in the Human Prefrontal Cortex. *Science*, *302*(5648), 1181–1185. https://doi.org/10.1126/science.1088545

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *Neuroimage*, *17*(2), 956–966.

Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and Language Syntax Interact in Broca's Area: An fMRI Study. *PloS One*, *10*(11), e0141069. https://doi.org/10.1371/journal.pone.0141069

Lashley, K. S. (1951). The problem of serial order in behavior. In *Cerebral mechanisms in behavior; the Hixon Symposium* (pp. 112–146). Wiley.

Leeuwenberg, E. L. (1969). Quantitative specification of information in sequential patterns. *Psychological Review*, *76*(2), 216.

Leggio, M. G., Tedesco, A. M., Chiricozzi, F. R., Clausi, S., Orsini, A., & Molinari, M. (2008). Cognitive sequencing impairment in patients with focal or atrophic cerebellar damage. *Brain: A Journal of Neurology*, *131*(Pt 5), 1332–1343. https://doi.org/10.1093/brain/awn040

Li, M., & Vitányi, P. (1993). An Introduction to Kolmogorov Complexity and Its Applications. In *An Introduction to Kolmogorov Complexity and Its Applications*. Springer New York. https://doi.org/10.1007/978-1-4757-3860-5

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, *4*(5), Article 5. https://doi.org/10.1038/87502

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *ELife*, *8*, e41541. https://doi.org/10.7554/eLife.41541

Maheu, M., Meyniel, F., & Dehaene, S. (2021). *Rational arbitration between statistics and rules in human sequence processing* (p. 2020.02.06.937706). bioRxiv. https://doi.org/10.1101/2020.02.06.937706

Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, *122*(3), 346–362. https://doi.org/10.1016/j.cognition.2011.11.003

May, P. J., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, *47*(1), 66–122. https://doi.org/10.1111/j.1469-8986.2009.00856.x

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houdé, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, *54*(3), 287–298. https://doi.org/10.1016/s0361-9230(00)00437-8

McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, *16*(4), Article 4. https://doi.org/10.1038/nn.3347

Meyniel, F., Maheu, M., & Dehaene, S. (2016). Human Inferences about Sequences: A Minimal Transition Probability Model. *PLOS Computational Biology*, *12*(12), e1005260. https://doi.org/10.1371/journal.pcbi.1005260

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Molinari, M., Chiricozzi, F. R., Clausi, S., Tedesco, A. M., De Lisa, M., & Leggio, M. G. (2008). Cerebellum and Detection of Sequences, from Perception to Cognition. *The Cerebellum*, *7*(4), 611–615. https://doi.org/10.1007/s12311-008-0060-x

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, *114*(42), E8913–E8921. https://doi.org/10.1073/pnas.1705373114

Moro, A. (1997). Dynamic Antisymmetry: Movement as a Symmetry-breaking Phenomenon. *Studia Linguistica*, *51*(1), 50–76. https://doi.org/10.1111/1467-9582.00017

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buchel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nat Neurosci*, *6*(7), 774–781.

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, *98*(2), 217–221. https://doi.org/10.1016/0304-3940(89)90513-2

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, *9*(11), Article 11. https://doi.org/10.1038/nrn2478

Nixon, P. D. (2003). The role of the cerebellum in preparing responses to predictable sensory events. *The Cerebellum*, *2*(2), 114. https://doi.org/10.1080/14734220309410

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, *88*(6), 1281–1296. https://doi.org/10.1016/j.neuron.2015.11.035

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), Article 7. https://doi.org/10.1038/nn1082

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140090. https://doi.org/10.1098/rstb.2014.0090

Planton, S., & Dehaene, S. (2021). Cerebral representation of sequence patterns across multiple presentation formats. *Cortex*, *145*, 13–36. https://doi.org/10.1016/j.cortex.2021.09.003

Planton, S., van Kerkoerle, T., Abbih, L., Maheu, M., Meyniel, F., Sigman, M., Wang, L., Figueira, S., Romano, S., & Dehaene, S. (2021). A theory of memory for binary sequences: Evidence for a mental compression algorithm in humans. *PLOS Computational Biology*, *17*(1), e1008598. https://doi.org/10.1371/journal.pcbi.1008598

Psotka, J. (1975). Simplicity, symmetry, and syntely: Stimulus measures of binary pattern structure. *Memory & Cognition*, *3*(4), 434–444. https://doi.org/10.3758/BF03212938

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, *77*(6), 481–495. https://doi.org/10.1037/h0029964

Restle, F., & Brown, E. R. (1970). Serial pattern learning: Pretraining of runs and trills. *Psychonomic Science*, *19*(6), 321–322.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*. https://doi.org/10.1126/science.274.5294.1926

Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research*, *152*(2), 229–242. https://doi.org/10.1007/s00221-003-1548-8

Santolin, C., & Saffran, J. R. (2018). Constraints on Statistical Learning Across Species. *Trends in Cognitive Sciences*, *22*(1), 52–63. https://doi.org/10.1016/j.tics.2017.10.003

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience*, *16*(4), Article 4. https://doi.org/10.1038/nn.3331

Schröger, E., Bendixen, A., Trujillo-Barreto, N. J., & Roeber, U. (2007). Processing of abstract rule violations in audition. *PloS One*, *2*(11), e1131. https://doi.org/10.1371/journal.pone.0001131

Shima, K., Isoda, M., Mushiake, H., & Tanji, J. (2007). Categorization of behavioural sequences in the prefrontal cortex. *Nature*, *445*(7125), 315–318. https://doi.org/10.1038/nature05470

Simon, H. A. (1972). Complexity and the representation of patterned sequences of symbols. *Psychological Review*, *79*(5), 369.

Simon, H. A., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review*, *70*(6), 534.

Southwell, R., & Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound sequences. *Cortex*, *109*, 92–103. https://doi.org/10.1016/j.cortex.2018.08.032

Summerfield, C., & de Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745–756. https://doi.org/10.1038/nrn3838

Taulu, S., Kajola, M., & Simola, J. (2004). Suppression of interference and artifacts by the signal space separation method. *Brain Topography*, *16*(4), 269–275. https://doi.org/10.1023/B:BRAT.0000032864.93890.f9

Todorovic, A., Ede, F. van, Maris, E., & Lange, F. P. de. (2011). Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *Journal of Neuroscience*, *31*(25), 9118–9123. https://doi.org/10.1523/JNEUROSCI.1425-11.2011

Todorovic, A., & Lange, F. P. de. (2012). Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. *Journal of Neuroscience*, *32*(39), 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: A whole-brain fMRI study. *NeuroImage*, *8*(1), 50–61. https://doi.org/10.1006/nimg.1998.0349

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*(5), 867–875. https://doi.org/10.3758/BF03193539

Uhrig, L., Dehaene, S., & Jarraya, B. (2014). A Hierarchy of Responses to Auditory Regularities in the Macaque Brain. *Journal of Neuroscience*, *34*(4), 1127–1132. https://doi.org/10.1523/JNEUROSCI.3165-13.2014

van Heijningen, C. A. A., de Visser, J., Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, *106*(48), 20538–20543. https://doi.org/10.1073/pnas.0908113106

Vitz, P. C., & Todd, T. C. (1969). A coded element model of the perceptual processing of sequential stimuli. *Psychological Review*, *76*(5), 433–449. https://doi.org/10.1037/h0028113

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751. https://doi.org/10.1038/nature02447

Wacongne, C., Changeux, J.-P., & Dehaene, S. (2012). A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. *Journal of Neuroscience*, *32*(11), 3665–3678. https://doi.org/10.1523/JNEUROSCI.5003-11.2012

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc Natl Acad Sci U S A*, *108*(51), 20754–20759. https://doi.org/10.1073/pnas.1117807108

Wang, L., Amalric, M., Fang, W., Jiang, X., Pallier, C., Figueira, S., Sigman, M., & Dehaene, S. (2019a). Representation of spatial sequences using nested rules in human prefrontal cortex. *NeuroImage*, *186*, 245–255. https://doi.org/10.1016/j.neuroimage.2018.10.061

Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Current Biology*, *25*(15), 1966–1974. https://doi.org/10.1016/j.cub.2015.06.035

Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved Sequence Processing in Primate Frontal Cortex. *Trends in Neurosciences*, *40*(2), 72–82. https://doi.org/10.1016/j.tins.2016.11.004

Wilson, B., Slater, H., Kikuchi, Y., Milne, A. E., Marslen-Wilson, W. D., Smith, K., & Petkov, C. I. (2013). Auditory Artificial Grammar Learning in Macaque and Marmoset Monkeys. *Journal of Neuroscience*, *33*(48), 18825–18835. https://doi.org/10.1523/JNEUROSCI.2414-13.2013

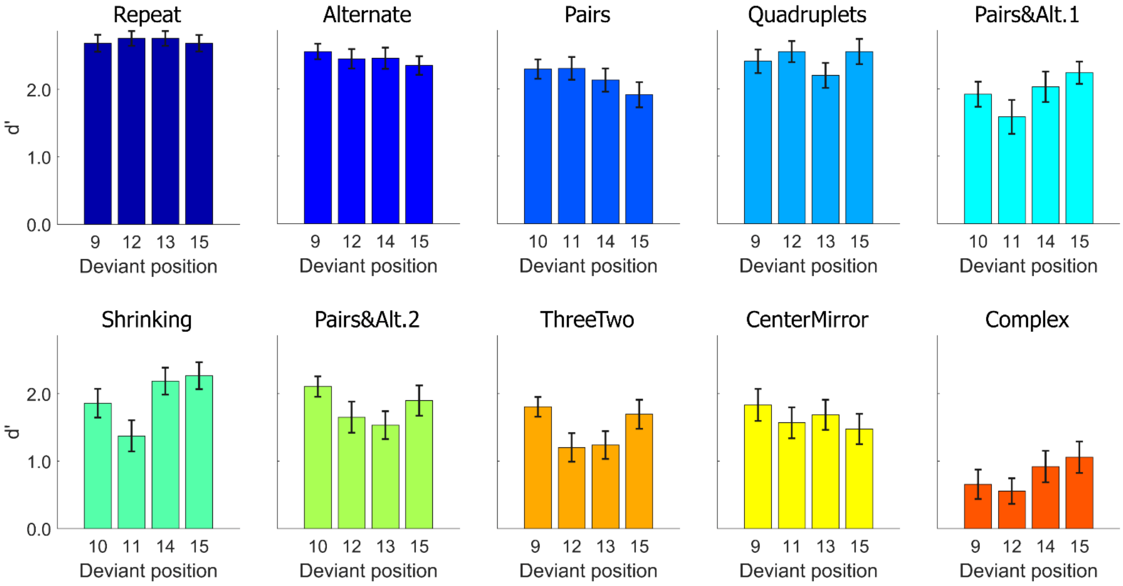


**Figure S1. Task performance: average sensitivity (d')**, for each position and each sequence. Error bars represent SEM.

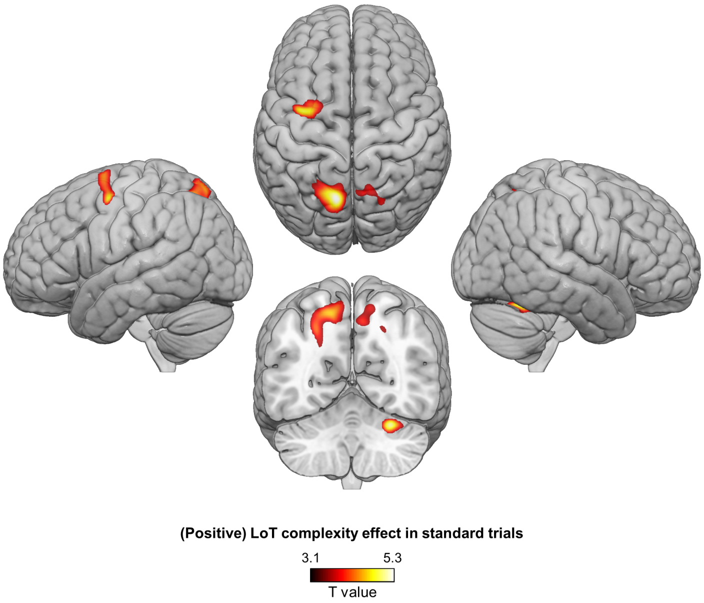


**Figure S2. Positive effects of LoT complexity effects on standard trials** (voxel-wise p < .001, uncorrected; cluster-wise p < .05, FDR corrected).

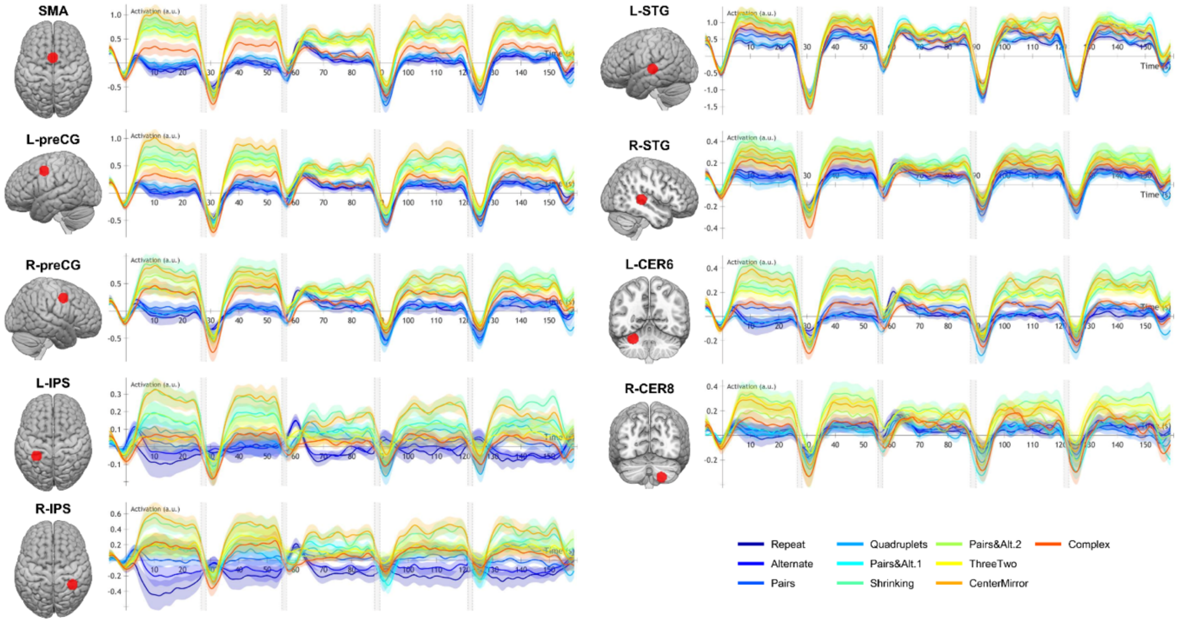


**Figure S3.** Time course of group-averaged BOLD signals for each sequence in nine ROIs where a LoT complexity effect was found. Each mini-session lasted 160-seconds and was composed of 5 blocks (2 habituation and 3 tests) interspersed with short rest periods of variable duration (depicted in light gray). The full time course was reconstituted by resynchronizing the data at the onset of each successive block (see Methods). Shading around each time course represents one SEM.

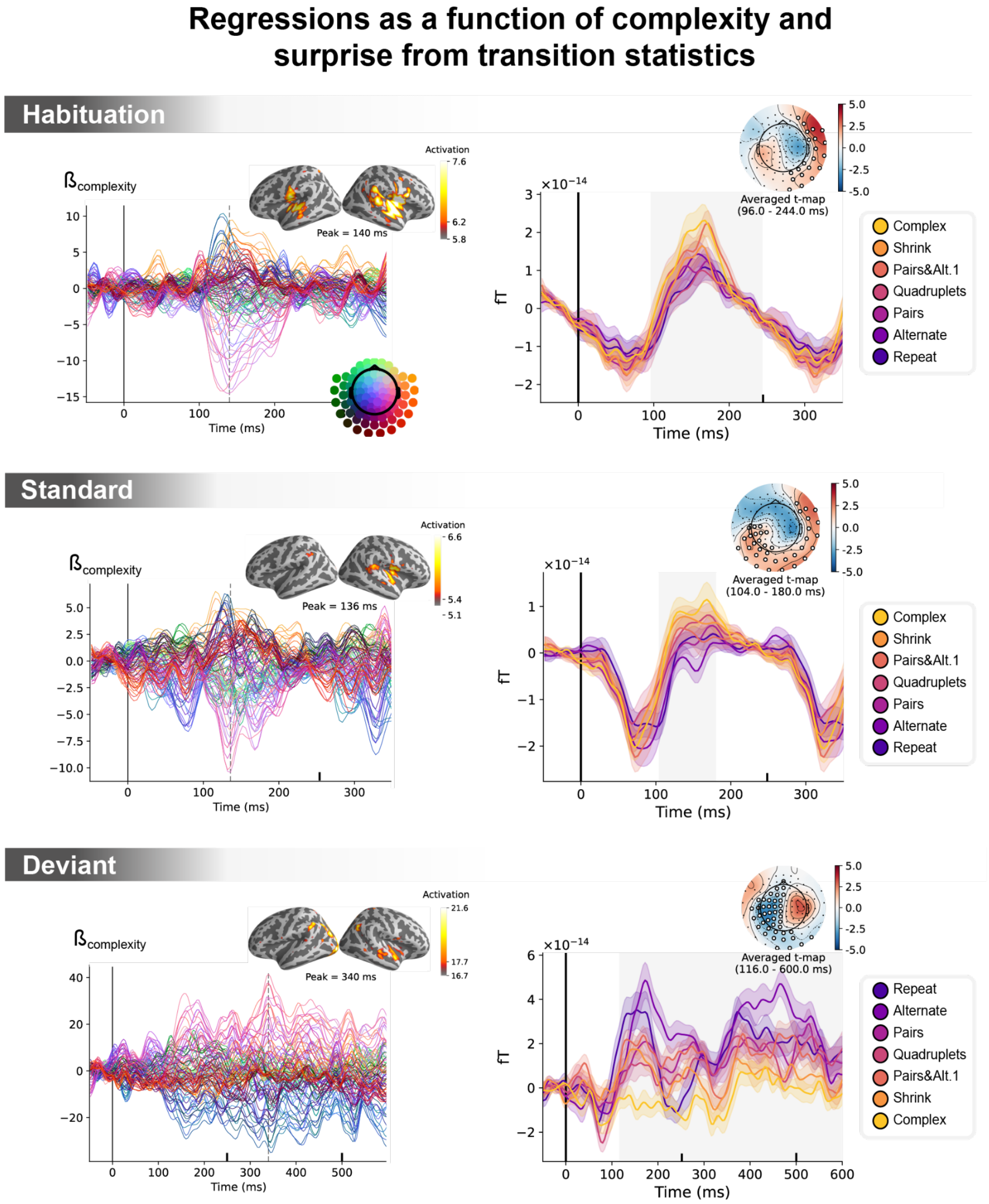


**Figure S4 Unconfounding the effects of statistical surprise and sequence complexity on MEG signals.** Left: amplitude of the regression coefficients ß of the complexity regressor for each MEG sensor, in a general linear model where surprise, repetition and alternation were also modeled. Insets show the projection of these coefficients on the source space for its maximal amplitude value, indicated by the vertical dotted lines. Right: illustration of spatiotemporal clusters where regression coefficients were significantly different from 0. Significant time windows are indicated by the shaded areas and have an opposite T-value for *Deviant* trials. Neural signals were averaged over the significant sensors for each sequence type and were plotted separately (see color legend). Note that the time-window goes until 600 ms for *Deviant* trials and until 350 ms for the other trials.

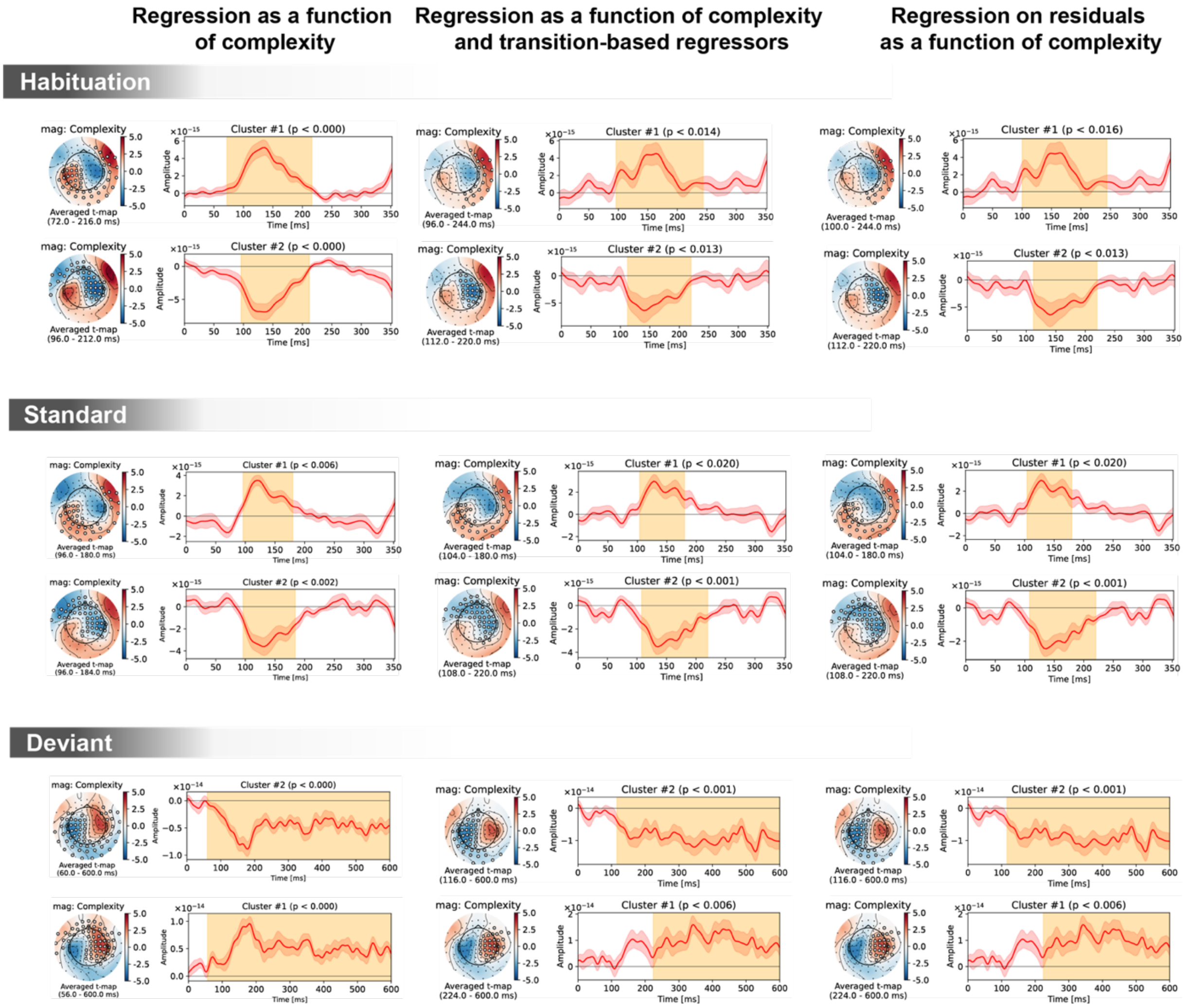


**Figure S5.** Significant spatiotemporal clusters for the complexity regressor in sensor space, shown separately for the 3 trial types (*Habituation, Standard, Deviant)* and 3 general linear models of MEG signals: with complexity alone (left column); with complexity, surprise and repeat/alternate (middle column); and with complexity after regressing out surprise and repeat/alternate signals. The clusters are very similar in all three cases, suggesting a robust effect of complexity irrespectively of transition statistics.

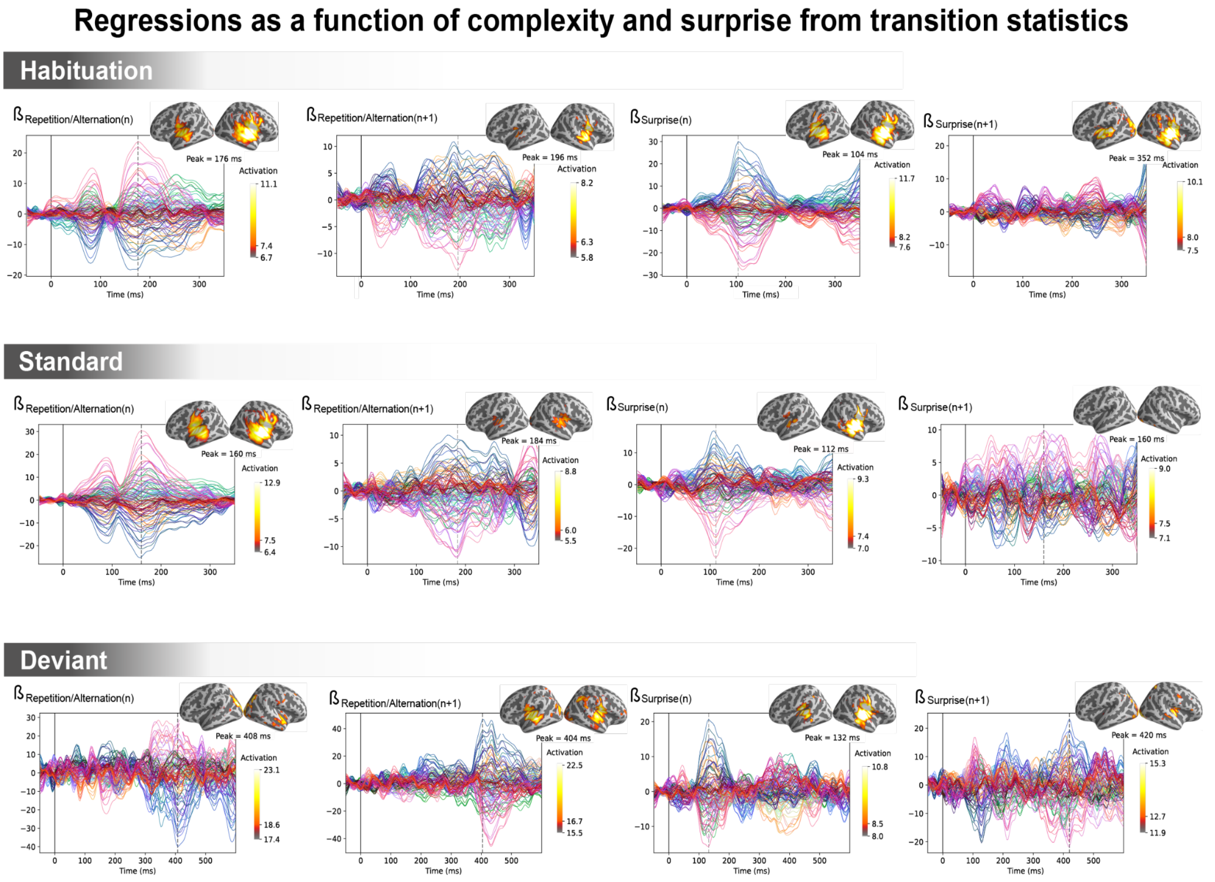
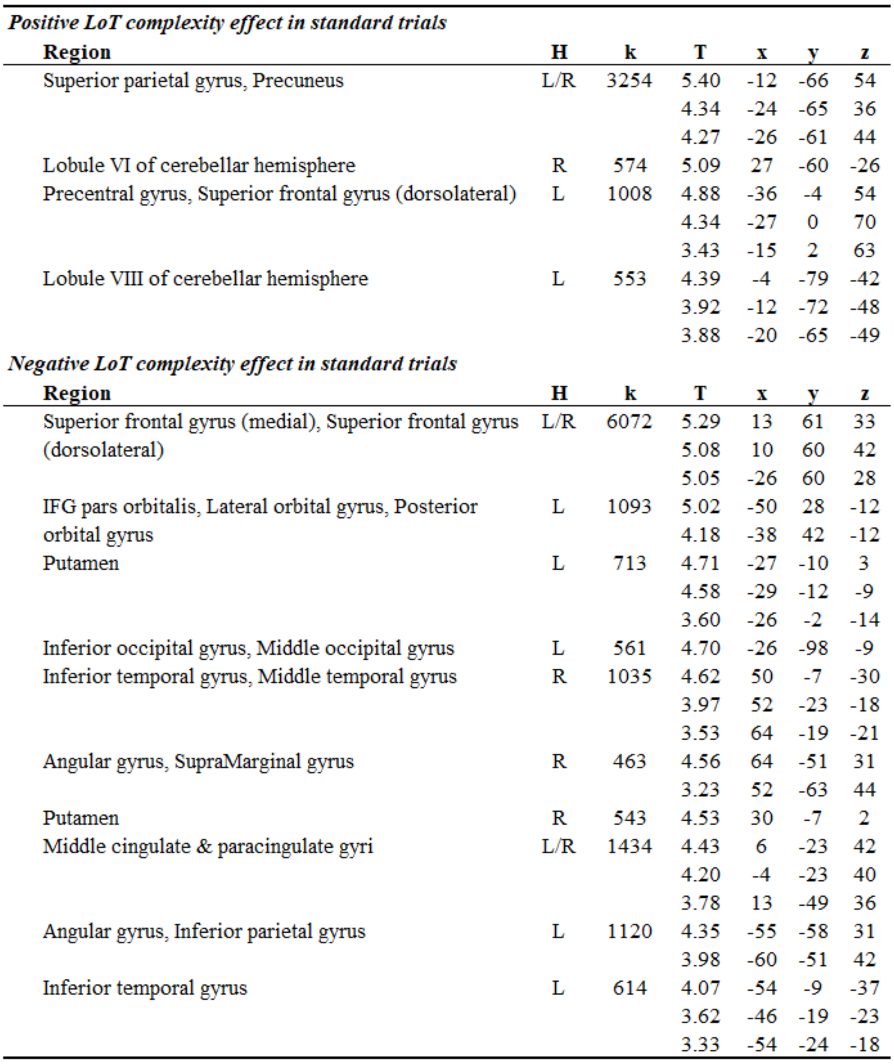


**Figure S6.** Amplitude of the regression coefficient ß for each MEG sensor for the 4 regressors of transition statistics: Repetition/Alternation for item *n* (presented at t=0 ms), Repetition/Alternation for item *n+1* (presented at t=250 ms), Surprise for item *n* and Surprise for item *n+1.* The Surprise predictor is computed using an ideal observer estimating surprise over 100 past observations. The projection on the source space at the time of its maximal amplitude is also shown.

**Table S1. fMRI complexity effect on standard trials** (voxel-wise p < .001, uncorrected; cluster-wise p < .05, FDR corrected)



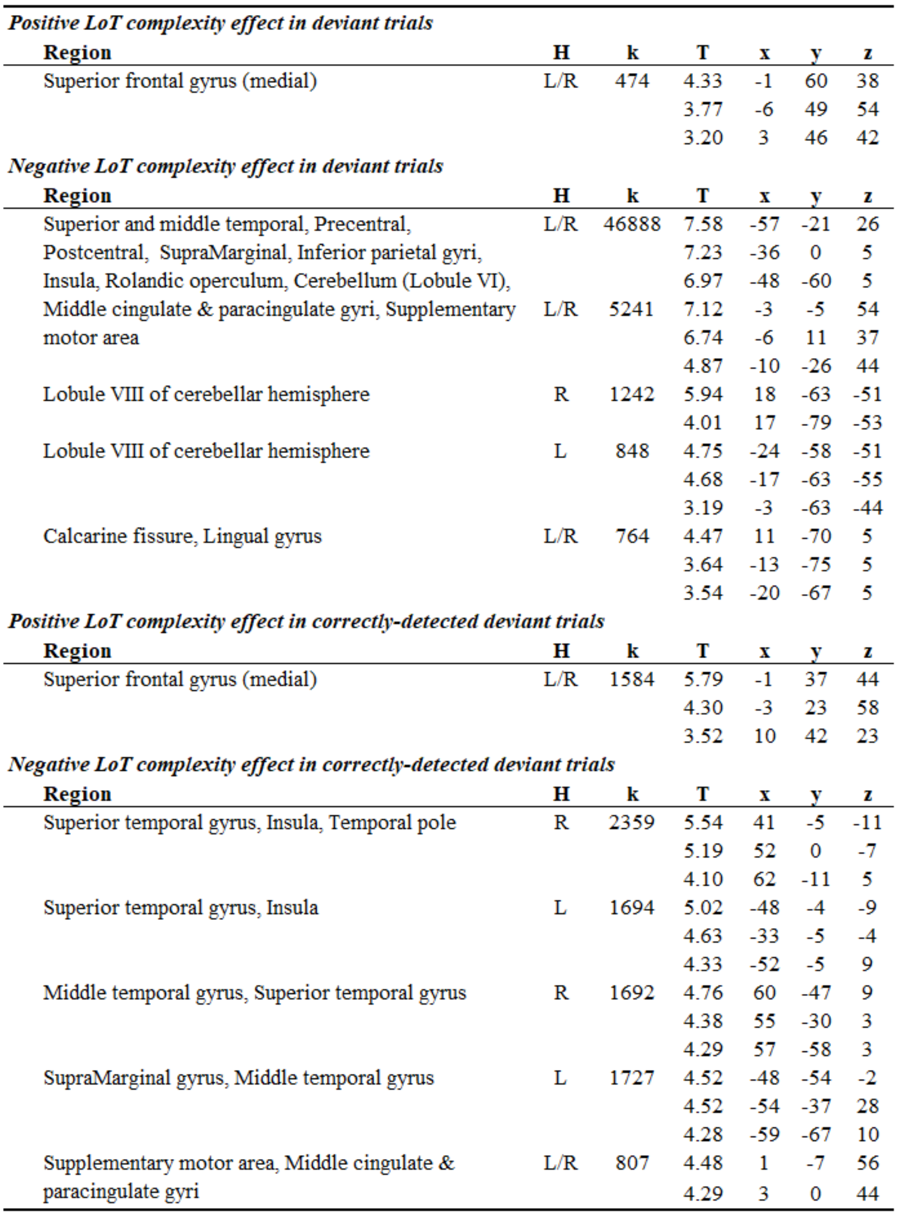
***Positive LoT complexity effect in standard trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior parietal gyrus, Precuneus | L/R | 3254 | 5.40 | -12 | -66 | 54 |
| | | | 4.34 | -24 | -65 | 36 |
| | | | 4.27 | -26 | -61 | 44 |
| Lobule VI of cerebellar hemisphere | R | 574 | 5.09 | 27 | -60 | -26 |
| Precentral gyrus, Superior frontal gyrus (dorsolateral) | L | 1008 | 4.88 | -36 | -4 | 54 |
| | | | 4.34 | -27 | 0 | 70 |
| | | | 3.43 | -15 | 2 | 63 |
| Lobule VIII of cerebellar hemisphere | L | 553 | 4.39 | -4 | -79 | -42 |
| | | | 3.92 | -12 | -72 | -48 |
| | | | 3.88 | -20 | -65 | -49 |

***Negative LoT complexity effect in standard trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior frontal gyrus (medial), Superior frontal gyrus (dorsolateral) | L/R | 6072 | 5.29 | 13 | 61 | 33 |
| | | | 5.08 | 10 | 60 | 42 |
| | | | 5.05 | -26 | 60 | 28 |
| IFG pars orbitalis, Lateral orbital gyrus, Posterior orbital gyrus | L | 1093 | 5.02 | -50 | 28 | -12 |
| | | | 4.18 | -38 | 42 | -12 |
| Putamen | L | 713 | 4.71 | -27 | -10 | 3 |
| | | | 4.58 | -29 | -12 | -9 |
| | | | 3.60 | -26 | -2 | -14 |
| Inferior occipital gyrus, Middle occipital gyrus | L | 561 | 4.70 | -26 | -98 | -9 |
| Inferior temporal gyrus, Middle temporal gyrus | R | 1035 | 4.62 | 50 | -7 | -30 |
| | | | 3.97 | 52 | -23 | -18 |
| | | | 3.53 | 64 | -19 | -21 |
| Angular gyrus, SupraMarginal gyrus | R | 463 | 4.56 | 64 | -51 | 31 |
| | | | 3.23 | 52 | -63 | 44 |
| Putamen | R | 543 | 4.53 | 30 | -7 | 2 |
| Middle cingulate & paracingulate gyri | L/R | 1434 | 4.43 | 6 | -23 | 42 |
| | | | 4.20 | -4 | -23 | 40 |
| | | | 3.78 | 13 | -49 | 36 |
| Angular gyrus, Inferior parietal gyrus | L | 1120 | 4.35 | -55 | -58 | 31 |
| | | | 3.98 | -60 | -51 | 42 |
| Inferior temporal gyrus | L | 614 | 4.07 | -54 | -9 | -37 |
| | | | 3.62 | -46 | -19 | -23 |
| | | | 3.33 | -54 | -24 | -18 |

**Table S2. fMRI complexity effect on deviant trials** (voxel-wise p < .001, uncorrected; cluster-wise p < .05, FDR corrected)



***Positive LoT complexity effect in deviant trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior frontal gyrus (medial) | L/R | 474 | 4.33 | -1 | 60 | 38 |
| | | | 3.77 | -6 | 49 | 54 |
| | | | 3.20 | 3 | 46 | 42 |

***Negative LoT complexity effect in deviant trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior and middle temporal, Precentral, Postcentral, SupraMarginal, Inferior parietal gyri, Insula, Rolandic operculum, Cerebellum (Lobule VI), | L/R | 46888 | 7.58 | -57 | -21 | 26 |
| | | | 7.23 | -36 | 0 | 5 |
| | | | 6.97 | -48 | -60 | 5 |
| Middle cingulate & paracingulate gyri, Supplementary motor area | L/R | 5241 | 7.12 | -3 | -5 | 54 |
| | | | 6.74 | -6 | 11 | 37 |
| | | | 4.87 | -10 | -26 | 44 |
| Lobule VIII of cerebellar hemisphere | R | 1242 | 5.94 | 18 | -63 | -51 |
| | | | 4.01 | 17 | -79 | -53 |
| Lobule VIII of cerebellar hemisphere | L | 848 | 4.75 | -24 | -58 | -51 |
| | | | 4.68 | -17 | -63 | -55 |
| | | | 3.19 | -3 | -63 | -44 |
| Calcarine fissure, Lingual gyrus | L/R | 764 | 4.47 | 11 | -70 | 5 |
| | | | 3.64 | -13 | -75 | 5 |
| | | | 3.54 | -20 | -67 | 5 |

***Positive LoT complexity effect in correctly-detected deviant trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior frontal gyrus (medial) | L/R | 1584 | 5.79 | -1 | 37 | 44 |
| | | | 4.30 | -3 | 23 | 58 |
| | | | 3.52 | 10 | 42 | 23 |

***Negative LoT complexity effect in correctly-detected deviant trials***

| Region | H | k | T | x | y | z |
|---|---|---|---|---|---|---|
| Superior temporal gyrus, Insula, Temporal pole | R | 2359 | 5.54 | 41 | -5 | -11 |
| | | | 5.19 | 52 | 0 | -7 |
| | | | 4.10 | 62 | -11 | 5 |
| Superior temporal gyrus, Insula | L | 1694 | 5.02 | -48 | -4 | -9 |
| | | | 4.63 | -33 | -5 | -4 |
| | | | 4.33 | -52 | -5 | 9 |
| Middle temporal gyrus, Superior temporal gyrus | R | 1692 | 4.76 | 60 | -47 | 9 |
| | | | 4.38 | 55 | -30 | 3 |
| | | | 4.29 | 57 | -58 | 3 |
| SupraMarginal gyrus, Middle temporal gyrus | L | 1727 | 4.52 | -48 | -54 | -2 |
| | | | 4.52 | -54 | -37 | 28 |
| | | | 4.28 | -59 | -67 | 10 |
| Supplementary motor area, Middle cingulate & paracingulate gyri | L/R | 807 | 4.48 | 1 | -7 | 56 |
| | | | 4.29 | 3 | 0 | 44 |