Cerebral representation of sequence patterns across multiple presentation formats

Samuel Planton, Stanislas Dehaene

Cerebral representation of sequence patterns across multiple presentation formats

Samuel Planton[a]*, Stanislas Dehaene[a,b]

[a] Cognitive Neuroimaging Unit, INSERM, CEA, CNRS, Université Paris-Saclay, NeuroSpin center, 91191 Gif/Yvette, France

[b] Collège de France, Université PSL Paris Sciences Lettres, Paris, France

* 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: samuel.planton@cea.fr

Abstract

The ability to detect the abstract pattern underlying a temporal sequence of events is crucial to many human activities, including language and mathematics, but its cortical correlates remain poorly understood. It is also unclear whether repeated exposure to the same sequence of sensory stimuli is sufficient to induce the encoding of an abstract amodal representation of the pattern. Using functional MRI, we probed the existence of such abstract codes for sequential patterns, their localization in the human brain, and their relation to existing language and math-responsive networks. We used a passive sequence violation paradigm, in which a given sequence is repeatedly presented before rare deviant sequences are introduced. We presented two binary patterns, AABB and ABAB, in four presentation formats, either visual or auditory, and either cued by the identity of the stimuli or by their spatial location. Regardless of the presentation format, a habituation to the repeated pattern and a response to pattern violations were seen in a set of inferior frontal, intraparietal and temporal areas. Within language areas, such pattern-violation responses were only found in the inferior frontal gyrus (IFG), whereas all math-responsive regions responded to pattern changes. Most of these regions also responded whenever the modality or the cue changed, suggesting a general sensitivity to violation detection. Thus, the representation of sequence patterns appears to be distributed, yet to include a core set of abstract amodal regions, particularly the IFG.

Keywords

Sequence processing; Language; Abstract pattern; Novelty detection; Functional magnetic resonance imaging

1. Introduction

Making predictions from past events is considered one of most fundamental function of the brain (Clark, 2013). When confronted with a temporal sequence of sensory elements, our brain constantly tries to anticipate what will be the next stimulus, thanks to the identification of the regularities governing the stimuli. Spontaneous learning of sequence regularities and associated expectations can be evidenced by observing the novelty responses triggered when these regularities are violated. The mismatch negativity (MMN), for instance, is observed when the repeated presentation of a "standard" sound stimulus is suddenly disrupted by the presentation of a different sound stimulus (Näätänen, 2003; Näätänen et al., 2007). Human sequence knowledge, however, goes way beyond the learning that a single stimulus is repeated with a specific timing, and extends to the encoding and manipulation of temporal sequences with a complex structure (Amalric et al., 2017; Bekinschtein et al., 2009; Planton et al., 2021; Wang et al., 2015). Sequence encoding and violation detection thus offers a simple paradigm in order to ask fundamental questions about the nature of the human grasp of abstract structures (Lashley, 1951), which underlies various higher-order human skills such as language, mathematics or music. Understanding when and how sequential structures are identified, encoded in the brain and used has emerged as a major issue for the elucidation of the origins of human uniqueness (Dehaene et al., 2015; Fitch & Hauser, 2004; Hauser et al., 2002).

A hierarchy of systems for sequence representation, with increasing degrees of abstraction, was proposed by (Dehaene et al., 2015), from the capacity to learn the transitions from one item to the next (as exemplified by the MMN paradigm) to the ability to manipulate nested tree structures, for instance the syntactic structures of language. Here, we focus specifically on the detection of what has been termed "algebraic patterns" (Marcus, 2019) , i.e. putative abstract templates underlying a sequence, such as AABB, independently of the specific identities of the items A and B. Pattern detection has been observed in infants (Kabdebon & Dehaene-Lambertz, 2019; Marcus et al., 1999), thus revealing a remarkably precocious ability for rule learning, which precedes and may underlie the subsequent emergence of language and mathematics (Marcus, 2019).

The local-global paradigm (Bekinschtein et al., 2009) was developed in order to dissociate the sensitivity to abstract patterns from lower-level responses to the transitions between sensory events. In this paradigm, two hierarchical levels are contrasted: a local one, related to the MMN (i.e., a sound differs from the preceding ones; AAAAB), and a global one, the level of patterns, when a whole sequence of stimuli, the deviant, differs from the preceding chunks, the standards (e.g. hearing AAAAA in the midst of many AAAAB sequences). A dissociation between these two levels was repeatedly

replicated : while the local effects is early (i.e., around 120 ms), mainly involves auditory sensory regions, and persists despite lack of consciousness, the global one is late (i.e. around 350 ms, the P3b EEG component), involves higher-order regions in the prefrontal and parietal cortex and is only observed in a conscious state (Bekinschtein et al., 2009; El Karoui et al., 2015; Strauss et al., 2015; Wacongne et al., 2011). The local-global paradigm is simple enough to be used with various human and non-human populations, without the need for specific task training (Basirat et al., 2014; Uhrig et al., 2014; Wang et al., 2015). Sequence processing studies based on such implicit learning and novelty response demonstrated that chunks of sound are easily encoded and spontaneously used for prediction, and found several candidate brain areas for pattern responses in temporal, prefrontal, and parietal cortices (Bekinschtein et al., 2009; Nourski et al., 2018; Strauss et al., 2015). Most crucially, Wang et al. (2015) showed that the global effect persists when, in the habituation phase, the pitch of the sounds, their duration, or their timing constantly varies from one sequence to the next. Such a novelty response implies the detection of a violation of an abstract pattern (e.g., AAAB or "three same, then one different"). Human areas sensitive to a change in the sequence pattern included the inferior frontal gyrus and the posterior superior temporal sulcus. Interestingly, those areas are also involved in the complex structures of language (Pallier et al., 2011), consistent with the idea that shared mechanisms are involved in sequence coding and high-level language processing. This study is just one in a long series of experiments (reviewed by Paavilainen, 2013) that examined to what extent the MMN violation paradigm involves abstract information processing, and revealed a sensitivity to abstract features such as pitch change (ascending or descending, Saarinen et al., 1992), pitch interval between sounds (Paavilainen et al., 1999), or melody contour independently of its pitch (Tervaniemi et al., 2001). Beside superior temporal and frontal regions, other authors have emphasized the role of the cerebellum in sequence processing (e.g. Molinari et al., 2008). Sensory prediction, was indeed proposed as a major function of the cerebellum in the context of motor learning (Nixon, 2003), but the integrity of the cerebellum also seems to be required to manipulate organized sequences of elements at a more abstract level (Leggio et al., 2008).

Nevertheless, several questions remain unsettled. Specifically, here we ask whether the abstract pattern underlying a sequence is automatically extracted whenever a sequence of stimuli is repeatedly perceived. Other experiments focusing on the extraction of abstract features have used a stable pattern, but introduced considerable variation in other physical dimensions of the stimuli (e.g. Ferrigno et al., 2020; Wang et al., 2015). In that case, the only regularity that can be inferred lies at the level of the abstract pattern, possibly associated with the mobilization of prefrontal regions such as the IFG. However, such abstract pattern encoding would perhaps not be systematically formed if a sequence was simply repeated in a fixed form and could therefore be encoded at a lower level, that of its sensory

features. It is unclear whether the representation of abstract patterns is an automatic by-product of sequence encoding, or whether it requires an active mobilization of a specific pattern-encoding system. Indeed, two versions of the hierarchical predictive coding theory may be contrasted: either the incoming sensory data is always captured at the lowest possible level at which a regularity occurs (and hence, in the present paradigm, we would not expect abstract pattern encoding); or, alternatively, abstract pattern extraction is an automatic process that applies even when other representational levels suffice to encode the input. The clarification of this issue is crucial for the interpretation of the results of the experimental paradigms used in several sequence processing studies (including the local-global paradigm, typically employing only two stimuli), especially when trying to make distinctions between the sequence processing skills of different populations (human vs. non-human primates, adults vs. prelinguistic infants). Furthermore, all of the abovementioned MMN and local-global studies used the auditory modality. They consequently often report the crucial involvement of the auditory-related cortex in the encoding of various forms of regularity, even for abstract features (Paavilainen, 2013). It is therefore unclear whether an independent representation of abstract patterns, detached from the sensory details, is actually identifiable in the brain.

Surprisingly, sequence processing experiments rarely manipulate the modality of stimulus presentation (although see Downar et al., 2000; Linden et al., 1999), thus potentially confounding the neural encoding of abstract and physical features of the stimuli. Only the first form of representation should generalize to a completely different set of stimuli, while the second would not. Interestingly, a recent study showed that sensory modality was a crucial determinant of cortical responses during working memory tasks with sequences, with different distinct prefrontal territories recruited for visual and for auditory stimuli (Michalka et al., 2015). Moreover, each of these prefrontal networks was linked to a specific information-processing domain, namely time for the auditory network and space for the visual network (Michalka et al., 2015). These results may suggest that, even in human prefrontal cortex, sequence representation remains tied to the input modality, and question the existence of regions dedicated to the automatic encoding of abstract patterns.

In the present study, we used fMRI to investigate how the neural response to pattern violations, in the context of passive processing of sequences, differ or are shared across different sequence presentation formats. We used two short binary algebraic patterns (i.e., AABB and ABAB) and four presentation formats, resulting from the combinations of a modality factor (visual vs. auditory) and a "cue" factor, distinguishing identity-based sequences, for which the two items are discriminated based on their physical properties (such as color or pitch), versus space-based sequences, for which the two items are discriminated based on their spatial location (left or right). Our main goal was to ask whether we could identify brain areas supporting a putative abstract representation of the pattern of

sequences, independently of the presentation format; or, alternatively, whether sequence patterns are essentially rooted in sensory, modality-specific representations of their constituent elements.

To demonstrate that a brain area contains an abstract code for an algebraic pattern, we suggest that several conditions must be met: (1) showing an habituation effect when confronted to the same pattern multiple times, characterized by a reduction of the BOLD signal over time, (2) showing a novelty response when the habituated pattern is violated, (3) doing so regardless of the modality or cue used to present the stimuli A and B; (4) showing a novelty response to pattern change even when the stimuli composing the sequence also change to a new cue or a new modality; (5) containing decodable information about the sequence pattern, again independently of the specific stimuli employed (as attested for instance by a similarity of brain activation patterns for all sequences supported by the same abstract algebraic pattern). Our experimental design allowed us to take advantage of several fMRI analyses techniques in order to seek for the neural representation of the different features of the sequences, including abstract algebraic patterns (see Davis & Poldrack, 2013). Univariate analyses helped us localize and track increases or decreases in activation during sequence habituation, as well as following an unexpected change in a specific sequence feature, in isolation or shared by multiple conditions (thanks to conjunction analyses). Additionally, multivariate analyses allowed us to search for regions that encoded abstract algebraic patterns, yet without any accompanying change in activation amplitude. The latter was achieved using by examining the similarity of activations patterns elicited by the different types of sequences (MVPA analysis).

A second objective of our experiment was to probe the specificity and the generalizability of the cerebral network previously found responsive to global sequence violations in the auditory local-global paradigm. Since global deviance effects were reported in prefrontal, temporal, parietal and cingulate regions (Bekinschtein et al., 2009; El Karoui et al., 2015; Nourski et al., 2018; Strauss et al., 2015; Wang et al., 2015), we expected to observe similar results with our “pitch-based” sequences (i.e. auditory-identity condition). We reasoned that the other three conditions with different sequences presentation formats (auditory-spatial, visual-identity, visual-spatial) would allow us to clarify, within this network, the areas involved in an amodal higher-order prediction system or “global workspace” network thought to underlie conscious processing (Dehaene & Changeux, 2011), and which are modality- or cue-specific. Furthermore, following Quirins et al. (2018) who showed that global deviance effects, in the auditory modality, were reflected in pupil dilation effects, we also tested on a subgroup of participants if similar results could be also obtained with the different sequences presentation formats. Pupil dilation, which was shown to be driven by several high-level cognitive processes (Hartmann & Fischer, 2014) was indeed proposed as a somatic marker of conscious access, and a counterpart of the P3b brain novelty response (Quirins et al., 2018; Zylberberg et al., 2012). We

also proposed here to test whether finer information (i.e. sensitivity to different types of novelty, including whether or not the sequence pattern changed) could be derived from such pupillometry analysis.

Finally, following the assumption that two human-specific high-level cognitive abilities, language and mathematics, are rooted in an elementary ability to process sequential patterns, a third objective of the experiment was to assess whether both language and math-responsive brain networks respond to abstract patterns and their violations, either fully independently of modality and sequence format, or perhaps tied to a preferred presentation, for instance language networks with auditory stimuli, and math networks with visuospatial stimuli (Michalka et al., 2015). Overlaps between non-linguistic sequence processing and the regions involved in mathematics and language processing have indeed been observed in the past (Dehaene et al., 2015; Wang et al., 2015, 2019). The identification of a shared substrate for higher-order cognitive functions and the encoding of minimal binary patterns might provide a key to understand the origins of human abstraction.

2. Materials & Methods

2.1. Participants

Twenty-two French young adults (mean age: 24.0, range 19-33; 10 women) took part in the study (a number chosen as a trade-off between scanning costs and the typical number of subjects in similar studies). All participants were right-handed, declared normal hearing and normal vision, and reported having no history of neurological nor language disorders. They received 80 euros for their participation and gave their written informed consent, conforming to the Helsinki declaration. The experiment was approved by the national ethical committee.

2.2. Stimuli

Four types of sequence trials were used in the experiment, manipulating two factors: the modality of the sequence (either visual or auditory), and the cue that distinguished the two items making up the sequence (either identity, e.g. two different pictures, two different sounds; or spatial, i.e. left vs right pictures and sounds). This resulted in four combinations: visual-identity, visual-spatial, auditory-identity, auditory-spatial; see Figure 1A). The stimuli were presented as sequences of four stimuli, using two possible sequence patterns: two-pairs (i.e. AABB) or alternation (i.e. ABAB).

The stimuli for the visual conditions were colored Gabor patches (Figure 1A, 2.9° diameter). The visual-identity condition used two stimuli: one red Gabor patch with 45° orientation, and one green

with 135° orientation. They were always presented in the center of the screen, on a dark grey background. The visual-spatial condition used a single stimulus, a yellow Gabor patch with 90° orientation, which could be presented either on the left or on the right of the central fixation cross (3.1° eccentricity).

The stimuli for the auditory conditions were complex tones synthesized with the superimposition of four sine waves, with frequencies chosen to correspond to musical notes. The auditory-identity condition used two sounds; a lower pitch sound (frequencies of 370Hz/554Hz/740Hz/1109Hz, corresponding to musical notes F#/C#/F#/C#) and a higher pitch sound (466Hz/698Hz/932Hz/1397Hz, notes Bb/F/Bb/F). The auditory-spatial condition used a single sound (415Hz/622Hz/831Hz/1245Hz, notes G#/D#/G#/D#), whose perceived location, left or right, was manipulated by reducing sound intensity by 50% in the opposite ear and by including an interaural time delay (ITD) of 600 µs (Babkoff et al., 2002). A fixation cross was always presented with the auditory stimuli.

2.3. Experimental design

16 different experimental runs were presented to each participant, each characterized by a different standard (dominant) sequence of stimuli. Each of the four trial types (see Figure 1A) was presented four times: two times using the pattern AABB (i.e. *aabb* and *bbaa*, using the same pair of stimuli but reversing their position in the sequence) and two times using the pattern ABAB (*abab* and *baba*).

Each experimental run was composed of four blocks of 20 trials (38 seconds) separated by rest periods (8 seconds). One trial (1000 ms) consisted in the presentation of the four items of the sequence (whether visual of auditory), each presented for 100 ms with a 150 ms blank (i.e. stimulus onset asynchrony, SOA, of 250 ms). Each trial was followed by an inter-trial period of variable duration (900 ms on average, with 150 ms standard deviation). In the first, habituation block, only the standard sequence was presented (20 times), while the three following blocks, the test blocks, contained 75% of standard trials and 25% of deviant trials. The 25% deviant trials were decomposed as follows: 5% in which the pair of stimuli remained the same but only the pattern changed (e.g. ABAB when the pattern of the standard sequence was AABB; “pure pattern change”), 5% in which only the modality of the trial changed (e.g. visual sequence when the standard sequence is auditory) while the cue and the pattern remained the same (i.e. “pure modality change”), 5% in which both the modality and the pattern changed, 5% in which only the sequence cue changed (i.e. “pure cue change”), and 5% in which both the sequence cue and the pattern changed (see Figure 1B).

The deviant trials were placed in the test blocks in a pseudo-random manner, i.e. with the constraint that the first two trials of each block could not be deviant trials and that each deviant trial was always followed by at least one standard trial. The same constraints applied to the selection of a subset of standard trials (5%), which were used in some of the analyses to allow a fairer comparison with deviant trials.

The 16 four-block experimental runs were presented in a different random order for each participant. For efficiency, they paired together to form 8 fMRI runs (approximately 6.5 minutes each) which contained two consecutive experimental runs with a different main trial type (separated by a 16 sec rest period).

The experimental paradigm required no overt response from the participant. However, to ensure that attention remained focused on the stimuli, questions were asked after each fMRI session. The participants were informed prior scanning that they would have to answer to some basic questions, but they did not know in advance which questions would be asked. Questions, selected depending on the trial type of the past session, could be as follow: “was there more auditory or visual stimuli?” (e.g., when the dominant trial type was auditory during the whole session), “which colors did you see?”, “did you noticed a dominant series of stimuli?”, “describe what you heard”, etc. Any participant who has remained attentive to the presented stimuli could easily answer each question correctly, and none of the questions involved counting or elaborate reasoning.

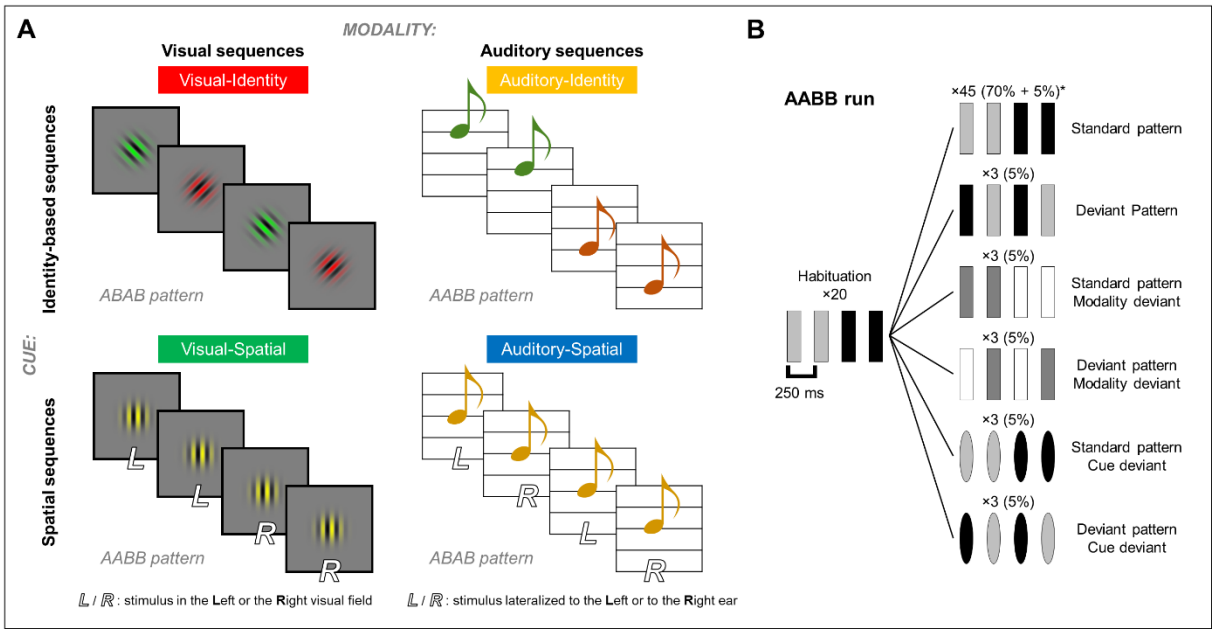


Figure 1: Experimental design and stimuli. A) Overview of the four different trial types, crossing the two experimental factors of modality (auditory or visual) and cue type (identity or spatial location). In each case, two sequence patterns, AABB or ABAB, were used, but the figure only shows a single example. Top-left: visual identity-based sequence, in which one stimulus is a red Gabor patch and the other a green Gabor patch (each presented in the center of the screen); top-right: auditory identity-based sequence, in which one stimulus is a low pitch tone and the other a high pitch tone; bottom-left: visual spatial sequence, in which one stimulus is a yellow Gabor patch presented on the left of the screen and the other a yellow Gabor patch presented on the right of the screen; bottom-right: auditory spatial sequence in which one stimulus is a tone with a left ear spatial dominance and the other the same tone with a right ear spatial dominance. **B)** Number (and proportion) of trials of each type. In each run, one sequence was designated as the frequent, standard pattern (75% of trials), interspersed with various types of deviants (25% of trials in total). This design is illustrated here in an experimental run in with standard pattern is AABB. Shades (light gray and black for one modality, dark gray and white for the other) and shapes (rectangles for one type of cue and ellipses for the other type) represent the different trial types shown in A). *In order to fully balance the design, 5% of the standard trials were treated in the same way as deviant trials, both in the pseudo-random generation of the trial list and in some of the analyses.

2.4. Localizer session

Besides the main experiment described above, all subjects also participated in 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), to which to new conditions were added: “rotated auditory” (meaningless auditory stimuli consisting in rotated sentences) and “false font” (meaningless visual stimuli of the same size and visual complexity as visual words). Other conditions of interest for the present purposes were “sentence reading”, “sentence listening”, “mental calculation visual” (mental processing of simple subtraction problems, such as 7 – 2, presented visually) and “mental calculation auditory” (the same but presented auditorily).

2.5. MRI data acquisition and preprocessing

MRI acquisition was performed on a 3T scanner (Siemens, Tim Trio), equipped with a 64-channel head coil. 208 functional scans covering the whole brain were acquired for each of the 8 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 mm^3, 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). Eye movements and pupil diameter size were collected using an EyeLink 1000 eyetracker (SR Research, ON, Canada). Due to technical difficulties related to the positioning of the participant in the scanner, eye-tracking signals of sufficient quality could only be obtained for half of the participants.

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, other regressors were computed in order to better correct for motion-related and physiological noise in the statistical models (using the PhysIO Toolbox, Kasper et al., 2017): 6 additional motion regressors corresponding to the derivatives of the first 6, and 6 regressors computed using the aCompCor method (Behzadi et al., 2007), applied to the CSF and to white matter (first 3 components of two principal component analyses). Additional regressors for motion outliers were also computed (framewise displacement larger than 0.5 mm; see Power et al., 2012), they represented 1.6% of volumes per subject on average.

Three participants were excluded from all analyses, two due to excessive head motion during acquisition (framewise displacement exceeding 3.0 mm on some sessions), one due to his inability to remain awake and attentive during acquisition (as observed via the eye tracker camera and as attested by his wrong answers to the basic questions following each session). Sample size for all fMRI analyses was thus n = 19.

2.6. fMRI analysis

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 128 s. Time series from the sequences of stimuli of each condition (each sequence modeled as an event of 850 ms duration) were convolved with the canonical hemodynamic response function (HRF). Two different first-level models were computed for each subject. Since we were interested in the cerebral substrates of sequence learning, in a first first-level GLM, the "habituation model", each 20-trials habituation block (from the 16 different experimental runs – each defined by a different standard trial type and pattern) was divided in five regressors, each representing the response to one of five successive bins of 4 trials (i.e. "part 1", "part 2", "part 3", "part 4" and "part 5"). The model also included, for each experimental run, one regressor for the standard trials of the test blocks and one regressor for the deviant trials (regardless of the deviant type). In a second first-level GLM, the "violation effects model", the five different deviant types (each representing 5% of test blocks) as well as a subset of 5% of standard trials (see Figure 1B), were modeled by a different regressor (in each experimental run). The model also included regressors for habituation trials and for remaining standard trials.

2.6.1. Univariate analyses

In order to estimate the effects of habituation to a given type of sequence, a first second-level within-subject ANOVA (group) analysis was conducted, using individual contrast images (beta maps) corresponding to the 20 different conditions of the "habituation model" (each of the four trial types × 5 parts of the habituation block). "Increasing activation" and "decreasing activation" T contrasts were computed for each trial type, as simple parametric modulations through the five bins (e.g. weights of [2 1 0 -1 -2] for "decreasing"). Second, in order to assess the main stimulus-dependent effects, a second-level within-subject ANOVA was conducted using 16 beta maps (from the "violation effects model") corresponding to the different types of standard (non-violated) trials. T tests were used to assess the effects of modality, cue and pattern. Finally, in order to assess the different violation effects, a third second-level within-subject ANOVA analysis was conducted using 24 beta maps (from the "violation effects model"), i.e. each of the four trial types × [5 deviant types + 1 standard]. T tests were used to assess the effects of each type of violation (in each trial type). Unless stated otherwise, all contrasts are reported with a voxel-wise threshold of $p < .001$ uncorrected and a cluster-wise threshold of $p < .05$ with FDR correction.

2.6.2. Multivariate analyses

As another way to investigate how the brain responded to the different dimensions of the sequences, we used a Representational Similarity Analysis (RSA) (Kriegeskorte et al., 2008). At the subject level, the main GLM was computed again but using non-normalized, non-smoothed volumes. RSA was conducted within each individual brain space using beta maps extracted from this GLM, using a spherical searchlight approach (Kriegeskorte et al., 2006). A 100 voxels searchlight was used (approximately 5 mm in radius).

For a first analysis focusing on the effects of the main condition factors in standard trials, we adopted a multiple regression RSA approach, with 3 model dissimilarity matrices (DSMs) representing the three factors of interest: modality, cue and pattern. The objective here was to identify putative brain areas sensitive to one factor (e.g., areas coding for the identity of the abstract pattern, AABB versus ABAB) while accounting for the two other factors (e.g. modality and cue). Since each factor had only two levels, the model DSMs were binary (e.g. for the modality DSM: 0 for pairs of conditions with the same modality, 1 otherwise; model DSMs are represented in Figure S1B). 16 volumes were used to compute the neural DSMs (one for each unique standard trial type). Spearman correlation was used as the distance metric. Beta maps for each regressor of the multiple regression RSA obtained for each individual were then normalized to the MNI space, and group-level statistical significance was assessed using one-sample t-tests (with voxel-wise threshold of $p < .001$ uncorrected, and cluster-wise threshold of $p < .05$ with FDR correction).

A second analysis was conducted with the deviant trials (but also including the 5% of control standard trials, see Figure 1B) using the same multiple regression RSA searchlight approach. Here, the objectives were to identify brain regions that encode a specific type of violation while accounting for the other types of violation and for the different dimensions of the trial type. Six (binary) model DSMs were used; 3 relative to the trial type (modality, cue, pattern), 3 relative to the type of violation (modality violation, cue violation, pattern violation; the 6 model DSMs are represented in Figure S2). 48 volumes were used, corresponding to the beta maps of three two-level factors for the trial type (modality × cue × pattern) × 6 deviants types (or rather, 5 deviants types + 1 standard). Since, due to the specifics of the experimental design, not all trial types were distributed equally in each block (e.g., visual modality-deviant trials with and without pattern violations necessarily fell in the same auditory blocks), leading to spurious correlations (and anticorrelations), cells of the dissimilarity matrix representing within-block correlations were excluded (10.6% of the total relevant matrix cells).

The two multiple regression RSAs were also performed using a region-of-interest (ROI) approach, using the three sets of ROIs described below (i.e. with each ROI deformed into each participant

individual space). All multivariate analyses were conducted using CoSMoMVPA (Oosterhof et al., 2016) and custom Matlab scripts.

2.6.3. ROI analyses

Different regions of interest (ROI) analyses were conducted, using individually and functionally defined ROIs. Three networks of regions were considered. Given that important connections of language processing abilities (especially syntax) and mathematics abilities with sequence processing abilities have been advocated in the literature, the first two sets of ROIs were regions involved in language and math processing.

The language-related network was composed of 7 regions of the left hemisphere, reported by 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) (Figure 8B). The mathematics-related network was composed of 7 regions, reported by 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) (Figure 8A). These exact two sets of regions were previously used elsewhere (Wang et al., 2019; although MFG was relabeled into preCG/IFG). Each of these language-related and mathematics-related ROIs was used as a mask in order to build functional ROIs (fROIs) for each participant, using the results of the independent localizer session. More specifically, for each area of the language-related network, we selected the 20% of most active voxels within the mask in the individual localizer contrast “Listening & reading sentences > Rotated auditory & false font”. Similarly, for each area of the mathematics-related network, we selected the 20% of most active voxels within the mask in the individual contrast “Mental calculation visual & auditory > Listening & reading sentences”. Subject-specific contrast value for each condition of interest in each fROI was then extracted.

In order to also probe areas involved in sequence learning, we defined a third network of ROIs, called the sequence habituation network, using data from the habituation blocks. Those ROIs were derived from the 2nd-level “decreasing activation” contrast map. More specifically, we kept voxels surviving a stringent threshold of $p < .05$ with FWE correction and within clusters larger 200 voxels, then used a watershed segmentation algorithm allowing the segmentation of large clusters containing distinct highly significant activation peak ($p < .001$, FWE corr.) into subclusters. 11 cortical ROIs were defined using this procedure: right inferior frontal gyrus / middle frontal gyrus (IFG/MFG), right inferior frontal gyrus / right precentral gyrus (IFG/preCG), left preCG (preCG), right posterior MFG (pMFG), posterior superior/middle temporal gyrus (pSTG/MTG), left IPS, right anterior IPS (aIPS), right posterior

IPS (pIPS), SMA, left and right anterior insula (aINS) (Figure 8C). We built the individual fROIs and extracted the contrast data using the same procedure as above, using the "decreasing activation" individual contrast from the habituation blocks as the localizer contrast.

Three sets of analyses were conducted in order to assess the response of each ROI to the different categories of deviant trials. First, we assessed if a novelty response could be observed. Second, we tested whether this novelty response differed depending on the type of violation. To these aims, we extracted the contrast estimates within each ROI and in each subject from the 12 contrasts of the different "pure change effects": pattern deviant > control standard (in each of the four main conditions), pure modality deviant > control standard (in each of the four main conditions) and pure cue deviant > control standard (in each of the four main conditions). For the third set of analyses, the 8 contrasts of the "pattern effects within modality or cue deviants" were considered: modality+pattern deviant > pure modality deviant (in each of the four main conditions) and cue+pattern deviant > pure cue deviant (in each of the four main conditions).

2.7. Pupillometry

Preprocessing of raw pupil diameter size timecourses during each session (2-ms sampling rate) included smoothing (30 ms moving average hanning window), removal of blinks and noise artifacts (identified as spikes in the derivative of the signal), interpolation of resulting missing segments (up to 500 ms, using piecewise cubic spline interpolation, which represented 5.0% of the data) and normalization (converted in percent change). 2500 ms epochs centered on stimulus onset (first stimulus of each trial) were then created for each trial, which included a 300 ms pre-stimulus baseline period (used for baseline correction). Only epochs with no remaining missing segments (following the artifact correction procedure) were considered good and kept. Reliable eye-tracker data was only available for a subpart of the participants. Only participants with more than 80% of good epochs were included in the pupillometry analyses, i.e. 11 participants (with 95.8% of good epochs on average, standard deviation = 4.8%). Contrasts between the average Event-Related Pupil Diameter (ERPD) of conditions of interest were then computed using non-parametric cluster-corrected permutation tests (5000 permutations) (see Quirins et al., 2018), using T-tests with a $p < .05$ threshold to identify clusters, and a $p < .05$ threshold to establish significance in the permutation test. All processing and statistical analyses were performed using custom Matlab scripts.

3. Results

3.1. Pupillometry

In the subgroup of participants for which reliable eye-tracking data was available (N = 11), we examined if the pupil gave behavioral evidence that the sequence violations were detected. A cluster-corrected permutation test revealed a significant increase in pupil size for deviant (including pattern violations, modality violations and cue violations) as opposed to standard trials ($p < .008$, starting 474 ms after first stimulus onset and maintained during the whole epoch). We then tested for pure pattern violation effects (e.g. AABB vs ABAB, without a modality or cue change). When pooling across all four block types, the pure pattern violation effect was significant (782-2200 ms temporal cluster, $p < .03$) (Figure 2A). When testing these effects within each block type (Figure 2B), the pure pattern violation effect was only significant during visual-spatial blocks (828-2200 ms, $p < .007$), although sub-threshold trends were also present in auditory-identity (762-1446 ms, $p = .064$), and auditory-spatial blocks (898-2200 ms, $p = .069$).

In order to further explore how the amplitude of the pattern violation effect interacted with the modality or the cue type, we computed an average "violation effect" in each participant and block type using a time window of interest. More specifically, we computed the average event-related pupil dilation (ERPD) in the [1016 – 2016 ms] time window (i.e. 1000 ms around the peak of maximal pupil dilation following the onset of a deviant trial) for each violation trial relative to all standard trials of the same block type. The violation effects were significant in visual-spatial (one-sample t-test: $t(10) = 2.77$, $p < .02$), auditory-identity ($t(10) = 2.28$, $p < .05$), and auditory-spatial ($t(10) = 2.84$, $p < .02$) conditions, but not in the visual-identity condition ($t(10) = 1.21$, $p = .29$) (Figure 2C). The violation effect amplitudes, however, did not differ across those four conditions: a linear mixed model with participants as the random factor, and block modality (two levels) and block cue (two levels) as fixed factors revealed no main effect of modality ($F(1, 494) = 1.92$, $p = .17$) nor cue ($F(1, 494) = 1.61$, $p = .21$), nor interaction between the two factors ($F(1, 494) = 0.002$, $p = .97$).

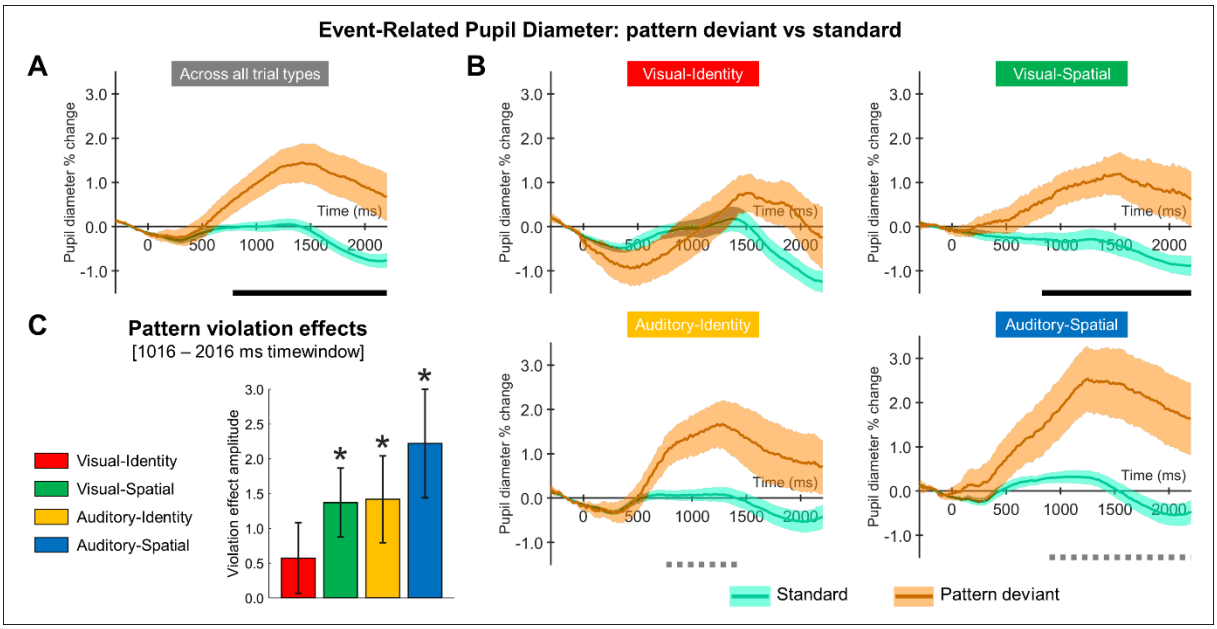


Figure 2: Behavioral evidence for deviant detection. Temporal evolution of the pupil diameter (expressed as % change relative to session average) on standard (green) vs. pure pattern deviant (red) trials, either averaged across all four block types (**A**), or separately for each of the four block types (**B**). Horizontal lines indicate significant temporal clusters, as identified by a cluster-corrected permutation test (solid black: $p < .05$, dashed gray: $p < .10$). Bands represent one standard error of the mean (SEM). **C**) Amplitude of pure violation effects in each of the four block types (computed using the average ERPD in the time window 1016-2016 ms). Error bars represent one SEM. *: $p < .05$ in a one-sample t-test.

We also assessed the detection of pattern violations (as seen from pupil diameter) when there was also a modality or a cue change. Thus, we compared modality+pattern or cue+pattern deviants, to pure modality or pure cue deviants. The main effect of an additional pattern violation was not significant, either when pooling modality and cue deviants in a single mixed model ($F(1, 1990) = 2.01$, $p = .16$) that also included block modality and block cue as factors; or separately for modality ($F(1, 977) = 0.15$, $p = .70$) and cue deviants ($F(1, 995) = 2.82$, $p = .09$). No significant clusters were found in cluster-corrected permutation contrasts, separately or when pooling across the four block types.

In summary, we observed some pupil dilation effects in responses to pattern deviant trials, which emerged around 500 ms after the onset of the first stimulus of the deviant sequence and peaked after around 1500 ms. Thus, the change in abstract sequence pattern (AABB versus ABAB) was detected and triggered a pupil dilation response. Using auditory sequences and the local-global paradigm, Quirins et al. (2018) proposed that pupil dilation may represent a somatic marker of conscious access, because (among other reasons) a pupil dilation in response to a violation was only observed when their participants were also able to report the presence of the regularity. Thus, the pupil results suggest that, although they were not explicitly responsive, participants detected the pattern change. Two caveats are in order: (1) the pupil response did not reach significant on visual-identity trial, perhaps due to the greater difficulty of those trials or to low power (11 participants, relatively small numbers

of deviant trials, and no significant interaction with modality or cue); (2) a pattern change did not induce any additional difference on pupil dilation, on trials where there was already a surprise due to modality or cue change. This may be an indication that, in such cases, the identity of the abstract pattern was not perceived or not attended to.

3.2. fMRI results

3.2.1. Habituation effects

We assumed that the cerebral regions involved in the encoding of the properties of a repeatedly presented sequence of stimuli would show an adaptation manifested by a decrease of activation over time. We thus searched for decreasing (or increasing) activations during habituation (see Materials & Methods). As shown in Figure 3A (see also Figure S1), a clear effect of the sequence modality was present: activity in bilateral visual related areas decreased on visual-identity and visual-spatial trials. A conjunction analysis (Figure 3B, left) indicated that a common set of occipito-temporal regions were involved (mainly IOG, FG, posterior lobe of the cerebellum) (note that the involvement of primary visual regions differed between the two conditions since the position of the stimuli differed). Conversely, activity in auditory related areas decreased on auditory-identity and auditory-spatial trials (bilateral Heschl's gyrus, superior and middle temporal gyrus, temporal pole), with the conjunction reaching significance in a right-dominant superior temporal area (Figure 3B, right). The conjunction between the two visual conditions also showed an habituation effect in bilateral anterior insula (aINS), bilateral precentral gyrus/middle frontal gyrus (preCG/MFG), bilateral inferior frontal gyrus (IFG), bilateral intraparietal sulcus (IPS) and bilateral supplementary motor area (SMA), while the conjunction between the two auditory conditions also showed an habituation effect in left posterior cerebellum, right aINS, right IFG and right IPS.

To identify brain areas putative involved in the learning of abstract patterns, independent of modality or cue, we also conducted a conjunction analysis of habituation effects over the four trial-type conditions (Figure 3C). Only two clusters reached the significance threshold: right IFG (pars opercularis, triangularis) and right IPS. As illustrated in Figure 3D for right IFG, their habituation profile was remarkably similar whether the sequence was visual or auditory, identity-based or space-based. With a slightly less strict cluster size threshold (k>50 voxels, but keeping the same voxel-wise threshold, justified by the fact that the conjunction null approach used here is quite conservative, especially since 4 contrasted are combined), 4 other regions also exhibited such a decreasing activation effect in all 4 conditions: left posterior cerebellum, SMA, left IPS and right aINS (see Table 1).

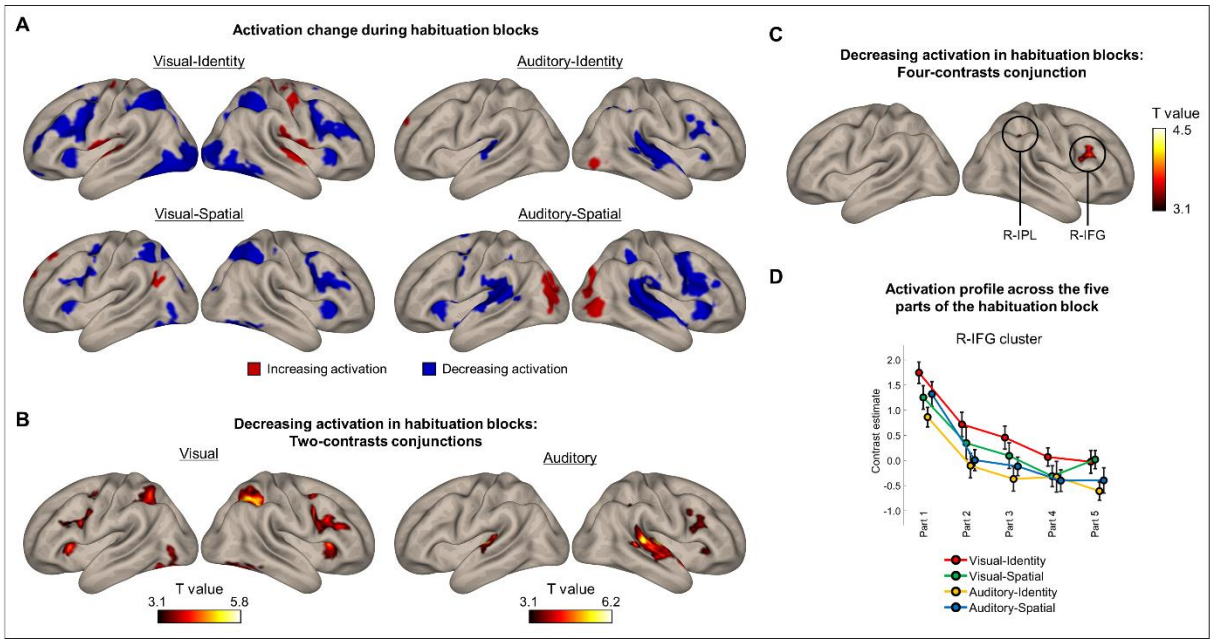


Figure 3: Habituation of the brain responses to the standard sequence. A) Habituation effect in each the four trial-type conditions (identified as increasing or decreasing responses across the five quintiles of the habituation block). **B**) Cue-independent habituation networks, identified by the conjunction of decreasing responses in the two visual conditions (left) and the two auditory conditions (right). **C**) Cue-independent and modality-independent habituation network (conjunction of the four trial types). **D**) Contrast estimates across the five quintiles of the habituation block, in each of the four trial-type conditions, for the right inferior frontal gyrus (IFG) cluster circled in C. Error bars represent S.E.M. All activation maps are reported with voxel-wise ($p < .001$, uncorrected) and cluster-wise ($p < .05$, FDR corrected) thresholds.

Table 1: Activation decrease in habituation: significant clusters in the conjunction over the four trial-type conditions (voxel-wise p<.001 unc., k>50)

Region	H	k	t	x	y	z
Inferior frontal gyrus (pars opercularis, pars triangularis)	R	496	4.54	46	16	31
			3.96	41	21	23
			3.21	50	12	17
Posterior cerebellum (lobule VI, crus I)	L	96	4.18	-31	-67	-27
			3.68	-26	-60	-30
Intraparietal sulcus (inferior parietal gyrus)	R	207	3.87	38	-46	38
			3.59	46	-37	49
Supplementary motor area	L/R	75	3.73	4	14	56
Intraparietal sulcus (inferior parietal gyrus)	L	68	3.67	-34	-42	40
			3.33	-41	-46	45
Anterior insula	R	52	3.51	34	26	0

3.2.2. Standard trials: main condition effects

We next examined, using both univariate and multivariate analysis, whether there were areas where activation varied as a function of sequence modality (visual vs. auditory), cue (identity vs.

spatial), as well as abstract pattern (AABB vs. ABAB). Such effects were assessed over all the standard trials of the experiment (including habituation); results are presented in Figure S1 and Table S2.

Unsurprisingly, univariate contrasts for modality (auditory > visual and visual > auditory) showed significant activation in bilateral auditory regions (STG, Heschl's gyrus) and bilateral visual regions (MOG, IOG, FG, lingual gyrus), as well as one cluster in the cuneus (for auditory > visual) and bilateral IPS clusters (for visual > auditory). The same regions were observed in the multivariate regression-RSA analysis, but with some additional areas (left and right superior preCG/MFG), indicating that, although their overall activation intensity did not differ, patterns in these areas encoded the sequence modality (Table S3).

Concerning the cue effect, in both univariate and multivariate analyses, it affected bilateral visual regions, mainly as a distinction between a bilateral pair of small clusters in the posterior lingual gyrus (spatial > identity) and a bilateral pair of large clusters in the MOG/IOG (identity > spatial). This is trivially related to the different positions of the visual stimuli on screen (see Figure 4A). Another pair of bilateral clusters was however found in the spatial > identity univariate contrast, in the posterior MTG/STG, which did not seem to be modality-dependent. Indeed, as illustrated in Figure 4A, the average beta value in these clusters was higher in auditory-spatial as opposed to auditory-identity trials, and higher in visual-spatial as opposed to visual-identity trials. This area, although a bit anterior to the middle temporal visual area (MT+) classically involved in visual motion processing, has already be reported as involved in auditory spatial and motion processing (Warren & Griffiths, 2003; Warren et al., 2002) as well as in the visual perception of biological motion (Pelphrey et al., 2005). It might thus reflect the fact that spatial sequences induced a percept of motion. Note however that this effect was not corroborated by the RSA analysis, since a sensitivity to the cue was only found in the bilateral occipital cortex.

Most crucially, in the whole-brain multiple regression RSA searchlight analysis, no areas showed a sensitivity to the sequence pattern (i.e. AABB vs. ABAB). However, a single cluster reached significance in the univariate contrast AABB > ABAB: right posterior cingulate/ventral precuneus (see Figure 4B). To explore further the activation in this area, beta values for each subject (and each of the 16 experimental runs) were entered in a linear mixed model with the factors "Modality", "Cue" and "Pattern" (with "Subject" as a random factor). Comparisons of least-squares means (with the *emmeans* R package) indicated that the activation in this area was stronger for the two-pairs (AABB) pattern as opposed to the alternating (ABAB) pattern in the conditions visual-identity (estimate: +0.44, $t(278) = 2.38$, $p < .018$), auditory-identity (+0.54, $t(278) = 2.93$, $p < .004$), auditory-spatial (+0.38, $t(278) = 2.06$,

p < .05), but not visual-spatial (+0.25, t(278) = 1.36, p = .18). It furthermore appeared quite consistent across blocks with the same pattern (i.e. starting with item *a* or item *b*, Figure 4B right).

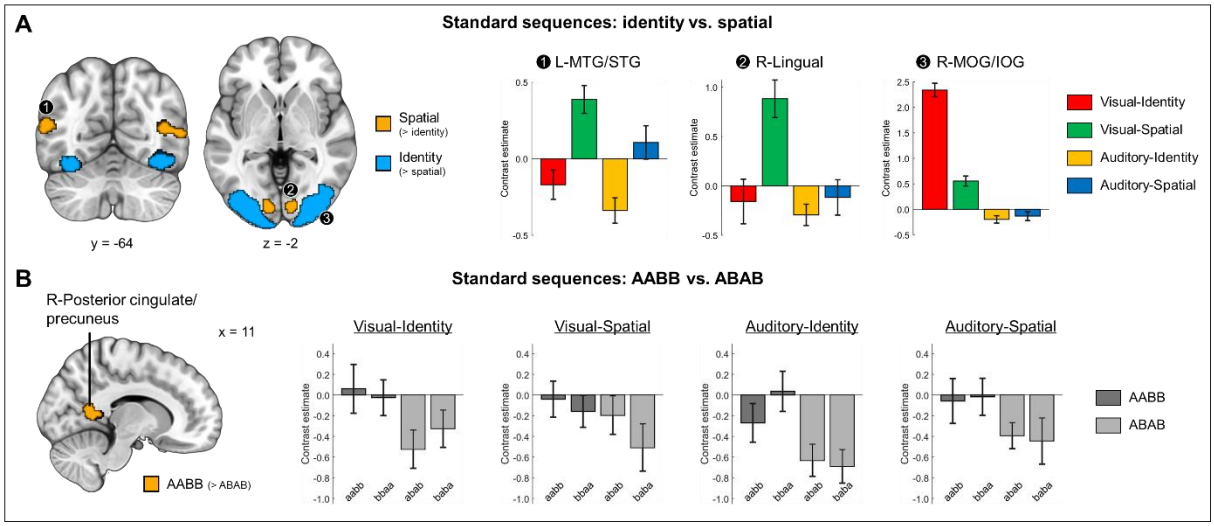


Figure 4: Brain regions distinguishing between the different types of sequences. A) Left: Significant activation clusters in the contrasts "Spatial > Identity" and "Identity > Spatial" restricted to standard trials (univariate analysis, voxel-wise p < .001 unc. and cluster-wise p < .05 FDR corr.). Right: Average contrasts estimates in each of three clusters of interest in each of the four block types. Cluster (1) shows a modality-independent effect. **B**) Left: Significant activation cluster in the contrasts "AABB > ABAB" performed with standard trials (univariate analysis, voxel-wise p < .001 unc. and cluster-wise p < .05 FDR corr.). Right: Contrasts estimates in this cluster for the four sequence orders (i.e. aabb and bbaa for the AABB pattern, and abab and baba for the ABAB pattern) in each of the four block types.

3.2.3. Violation effects

Univariate results

The participants were confronted to 3 types of "pure violation" trials, where either the pattern, or the modality, or the cue changed relative to the standard trials. In order to investigate how the brain responded to such deviants, pure violation trials were contrasted against a subset of standard trials (5%, pseudo-randomly selected from all standard trials, see Materials & Methods).

A large bilateral network responded to a violation of the standard pattern (i.e. pure pattern change effect): see Table 2 and Figure 5A. In order to better delineate the most important nodes of this network a very stringent threshold was used in a second stage (voxel-wise threshold of p < .001 with FWE correction; Figure 5A). Some activation survived in the SMA and right IPS, but also (smaller clusters) in the right STG, right preCG/MFG, left aINS/IFG and left IFG (pars opercularis). Note that the regions reported above as involved in a domain-general habituation to sequences, clearly overlap with this pure pattern violation network. Some of them, especially IFG, were previously found to respond to pattern violations in the auditory modality (Bekinschtein et al., 2009; Wang et al., 2015).

Table 2: Significant activation clusters in the "pattern deviant > (control) standard" contrast (voxel-wise p < .001 unc. and cluster-wise p < .05, FDR-corr)

Region	H	k	t	x	y	z
Supplementary motor area	L/R	1918	7.68	-3	7	63
			7.55	-5	12	52
			5.65	6	14	54
Superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus	R	3074	7.57	46	-23	-6
			5.98	46	-37	3
			5.98	55	-39	9
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	R	3555	7.25	46	11	30
			6.28	38	4	37
			6.18	46	7	44
Intraparietal sulcus (inferior parietal gyrus), postcentral gyrus, supramarginal gyrus	R	2042	7.17	48	-33	51
			6.45	41	-39	42
			5.15	29	-54	44
Insula, inferior frontal gyrus (pars triangularis, pars orbitalis), temporal pole	L	780	6.20	-33	30	3
			4.01	-57	4	-7
			4.01	-52	12	-9
Inferior frontal gyrus (pars opercularis, pars triangularis)	L	2022	6.18	-48	7	17
			5.87	-40	0	52
			5.57	-26	-5	54
Middle temporal gyrus, superior temporal gyrus	L	1445	5.95	-48	-42	24
			5.90	-64	-23	0
			5.71	-50	-49	7
Intraparietal sulcus (inferior parietal gyrus), postcentral gyrus	L	851	5.94	-41	-39	44
			4.07	-27	-39	40
Insula, inferior frontal gyrus (pars triangularis)	R	520	5.31	36	25	0
			4.12	52	35	3
Cerebellum (lobule VIII, lobule VIIB)	L	556	5.16	-15	-77	-48
			5.10	-27	-67	-55
Cerebellum (lobule VI, crus I)	R	200	5.12	39	-63	-27
Temporal pole, superior temporal gyrus	R	237	5.02	57	9	-11
			4.20	46	0	-13
Cerebellum (lobule VI, crus I)	L	289	4.60	-34	-70	-27
			4.59	-19	-68	-27
			3.71	-31	-63	-28

Was pattern change detected in the same regions regardless of modality or cue? Within each of the main four block types, we observed both common and condition-dependent responses to pattern change (Figure 5B, left). The occurrence of a pattern deviant in visual-spatial and visual-identity conditions caused an increased activation in the same set of bilateral areas (IFG, preCG, IPS, aINS, SMA, cerebellum), including a right inferotemporal cluster. Additionally, a bilateral posterior MTG/STG cluster was present in the visual-spatial but not the visual-identity condition (at the same location of the area previously found as more responsive to spatial sequences, Figure 4A) and a larger and more intense activation cluster was observed in the SMA in the visual-identity condition. In the auditory modality, no inferotemporal cluster was observed, but rather the involvement of the bilateral STG and MTG (together with SMA and a right-dominant IFG). Activation was wider and more intense in the

bilateral STG/MTG in the auditory-identity condition, and conversely in the bilateral IFG for the auditory-spatial condition. Auditory and visual conditions also differed in the parietal activation, which was stronger and wider (with a right-hemisphere dominance) in the visual conditions, and absent in the auditory conditions (with the exception of a small cluster in the auditory-identity condition). Although the main effect yielded highly significant results, at the predefined threshold no cluster survived in a conjunction between the four contrasts (i.e. pure pattern violation effect in each condition). Here again, since the predefined statistical threshold was quite conservative for a conjunction null approach with four contrasts, a slightly lower threshold was used in a second stage ($p < .005$ unc., $k > 40$). Three clusters were found (Table 3): right IFG, right STG (in the depth of the superior temporal sulcus) and SMA. This reinforces the assumption these three areas play an important role in the detection of pattern violations in a modality-independent manner.

Table 3: Significant activation clusters in the conjunction over the four trial-type conditions of the "pattern deviant > (control) standard" contrast (relaxed threshold: voxel-wise $p<.005$ unc., $k>40$)

Region	H	k	t	x	y	z
IFG (pars opercularis, pars triangularis)	R	143	3.78	50	18	26
Superior temporal gyrus	R	40	3.54	46	-23	-6
Supplementary motor area	L	106	3.15	-3	7	63
			3.13	-5	12	52
			2.78	-1	4	72

Although all these areas are potential candidates for the encoding of the pattern, as they respond to a pure pattern change, the novelty response they exhibit is not necessarily restricted to this type of violation. Indeed, very similar networks were observed for the main effect of pure modality violation (Figure 5A, middle, and Table S4) and the main effect of pure cue violation (Figure 5A, right, and Table S4) (i.e. when the stimuli are different but the abstract pattern remains the same). The modality violation effect was however characterized by a strong activation of a bilateral posterior STG/temporo-parietal junction (TPJ) cluster, which was less intense (and more right-lateralized) in the two other violation effects. The IPS activations were larger and more anterior for the cue violation effect than for the other two. In contrast with the pattern violation effect, both modality violation and cue violations effects included cingulate gyrus and precuneus clusters. Regarding effects within each condition (Figure 5B, middle and right), we observed an increased activation of networks related to the modality and cue of the deviant (e.g. visual regions activate to an unexpected visual sequence within an auditory block, superior/middle temporal regions activate to an auditory-identity sequence amidst auditory-spatial ones or conversely). Here again, IFG and IPS regions show the most consistent response to the presence of any type of violation. The modality effects are consistent with previous work showing unimodal responses to changes in the sensory environment in sensory-related areas, and multimodal

ones in right TPJ, IFG, SMA and aINS (Downar et al., 2000). The cue violation effects confirms the importance of posterior MTG and IPS regions in multimodal spatial localization and motion processing (Bushara et al., 1999; Lewis et al., 2000; Warren & Griffiths, 2003).

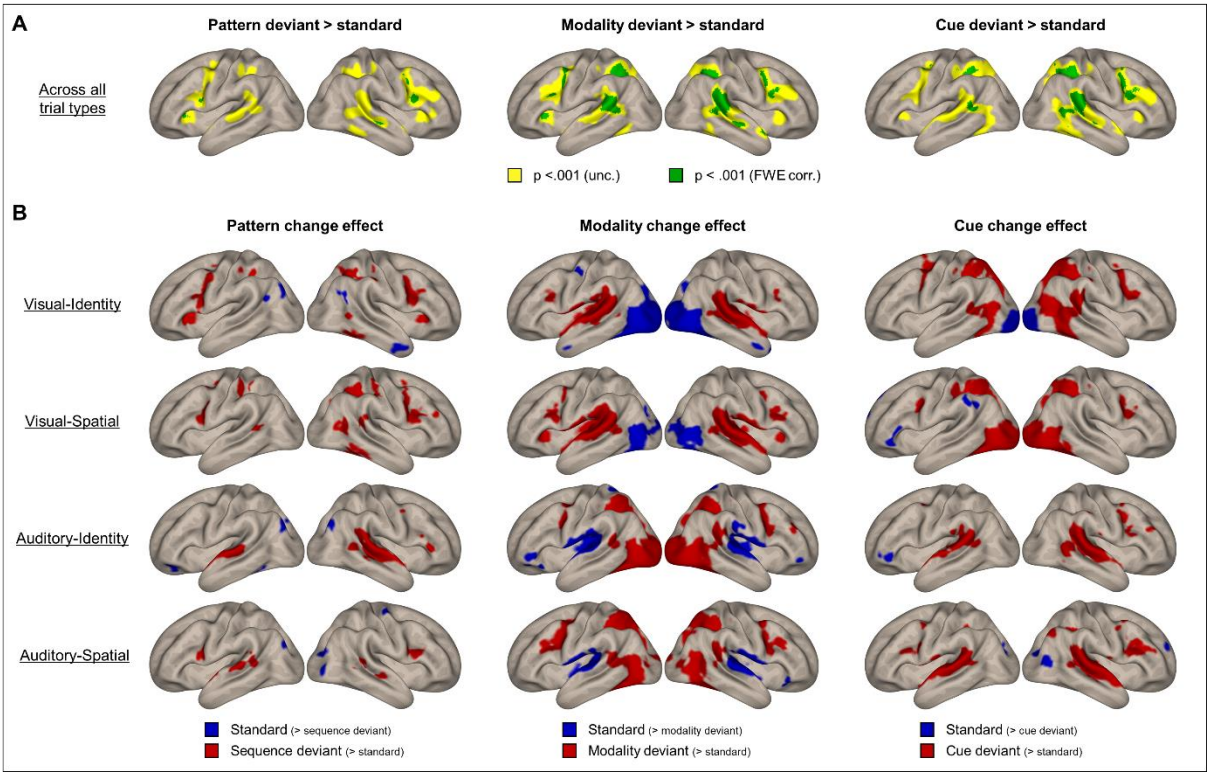


Figure 5: Brain regions responsive to a violation of the habitual sequence. A) Three main (pure) violation effects. From left to right: pattern deviant (i.e. pure pattern change relative to the pattern of the standard trial of the block), modality deviant (i.e. pure modality change) and cue deviant (i.e. pure cue change), each contrasted with standard sequences (i.e. the controlled subset of 5% of standard trials). Results are shown with two different voxel-wise thresholds: $p < .001$ with no correction in yellow and $p < .001$ with FWE correction in green (each also with cluster-wise threshold of $p < .05$ FDR corrected). **B**) Three violation effects (pure pattern change, pure modality change, pure cue change) computed separately for each trial type (i.e. trial type of the standard sequence of the block). Standard > deviant contrast is shown in red, deviant > standard is shown in blue (voxel-wise $p < .001$ uncorrected and cluster-wise $p < .05$ FDR corrected).

Although the three types of violations shared largely similar network, three areas showed a larger response to pattern violations than to modality and cue violations combined (contrast *+2 -1 -1* computed across all conditions): dorsal part of the preCG/SFG, pars orbitalis of right IFG and pars orbitalis/triangularis of left IFG (Figure 6).

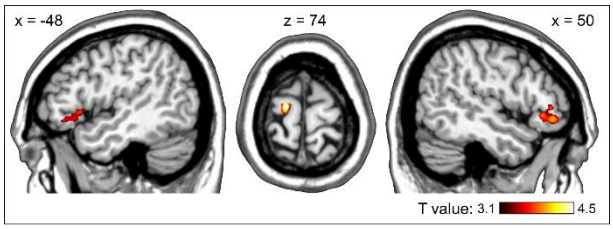


Figure 6: Brain response specific to pattern deviants, as identified by the contrast *Pattern deviants > [Modality & Cue deviants]* contrast (voxel-wise p < .001 uncorrected and cluster-wise p < .05 FDR corrected).

As a next step in our search for brain regions coding for abstract patterns, we next examined whether, when the stimuli changed (cue or modality violation), some brain regions showed an additional sensitivity to the fact that the abstract sequence pattern (AABB or ABAB) also changed – a marker of cross-modal or cross-cue generalization. To this end, we examined the contrasts for "modality+pattern violation vs. pure modality violation" and "cue+pattern violation vs. pure cue violation". No significant activation (in either direction) was found for the modality-based contrast. For the cue-based contrast, activation in a left cerebellum cluster (posterior lobe, Crus I, Crus II) was higher when the pattern was the same as opposed as when it was different. Note that since activation in this cluster was below the baseline level, it can also be interpreted as an increased deactivation when the pattern was additionally violated. We then looked at these two contrasts, in the positive and negative direction, within each of the four trial types. Only two of them yielded significant activation, both during the visual-spatial condition. One anterior cingulate cluster showed an increased deactivation for a modality+pattern violation as opposed to a pure modality violation (i.e. when the deviant trial was Auditory-Spatial) (k=481; x=8, y=39, z=17; T=4.93), and one SMA/paracentral lobule cluster showed an increased activation for a cue+pattern violation as opposed to a pure cue violation (i.e. when the deviant trial was visual-identity) (k=523; x=8, y=-19, z=58; T=4.35). Note that this cluster was more posterior than the SMA cluster observed for the cue violation effect (y=-19 against y=12).

In summary, in agreement with the pupil dilation results, fMRI indicated that abstract pattern generalization effects were overall very limited and absent in most conditions. When a violation of modality or cue occurred, activation shot up in many areas, but there was no additional signal change when the sequence pattern was also violated.

Multivariate results

The multiple regression RSA searchlight analysis for deviant trials was conducted by examining, within each subject, 48 fMRI images representing the average brain responses to the 5 different sets of deviant trials + 1 set of control standard trials), and modelling their 48x48 dissimilarity matrices with 6 theoretical dissimilarity matrices (DSM): "trial modality", "trial cue", "trial pattern", "modality violation", "cue violation" and "pattern violation". Note that we did not considered the type of the

block in which a given trial was presented but the type of the deviant trial. This allows to better disentangle between areas sensitive to a specific sensory property, from those sensitive to a change in the specific sensory property. The regression resulted in 6 whole-brain maps per participant representing the 6 betas of the regression. In this way, we could test whether, for instance, the activation patterns in some brain areas encoded the presence or the absence of a violation of the sequence pattern, whether or not there was also a modality or a cue violation, and regardless of the modality/cue/pattern of the trial. Significance was assessed using one-sample T-tests with voxel-wise threshold of $p < .001$ (uncorrected) and cluster-wise threshold of $p < .05$ (with FDR correction). Together with this main analysis with all 6 trial types, two secondary analyses were conducted using either the two 2 types without cue nor modality violation (i.e. testing for a pure pattern violation effect against standards), or the 4 trial types with modality or cue violation (i.e. testing pattern and pattern violation effects when the set of stimuli also changed, as in the above univariate analysis).

Regarding the type of the trial, replicating the analysis previously conducted with standard trials, significant effects were found for the modality and the cue DSMs but not for the pattern DSM. The networks of areas showing a sensitivity to the modality or the cue of the deviant trial were very similar to the ones found with standard trials: superior temporal regions, occipital regions, IPS, preCG/MFG and IFG for the modality DSM, occipital regions for the cue DSM (see Figure 7A).The modality sensitive network was however much more extensive (and with stronger effect sizes) than in the univariate analysis or in the RSA with standard trials. The fact that modality was more easily decoded on deviant trials may be due to a stronger activation to the same sequence when it is novel as opposed to when it has already been presented multiple times (due to habituation). Regarding the type of violation, six large clusters were obtained for the modality violation DSM, located in left temporo-parietal cortex, right temporo-parietal cortex, left preCG/IFG, right preCG and right cerebellum (Table 4). These regions thus encoded the information about whether or not the modality of the unexpected deviant trial was the same as the one of the standard trial of the block (regardless of all other trial properties including its actual modality). They partly correspond to the set of areas reported to respond to modality violations in the univariate analyses (with the exception of the bilateral aINS). The cue violation effect, traducing the ability to detect when an identity-based sequence is replaced by a space-based one (and conversely) was much more restricted, with one cluster in the IPS/supramarginal gyrus ($k=161$; $x=43$, $y=-32$, $z=38$; $T=5.58$).

Table 4: Significant activation clusters for the modality violation DSM in the multiple regression RSA searchlight analysis with deviant trials (voxel-wise $p < .001$ unc. and cluster-wise $p < .05$, FDR-corr)

Region	H	k	t	x	y	z
Superior temporal gyrus, middle temporal gyrus, supramarginal gyrus, precuneus, inferior parietal gyrus, superior parietal gyrus	L	7346	10.07	-59	-33	26
			9.14	-55	-40	28
			8.88	-59	-46	12
Superior temporal gyrus, middle temporal gyrus, supramarginal gyrus, temporal pole	R	7069	9.27	53	-40	21
			9.22	62	-25	2
			8.84	53	12	-7
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus	L	3158	8.06	-41	0	54
			6.52	-45	11	26
			6.40	-48	2	38
Supplementary motor area, middle cingulate gyrus	L/R	932	6.15	-3	5	37
			5.68	-3	9	47
			5.53	-5	0	23
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	R	615	5.87	34	0	44
			4.99	48	5	45
			4.80	50	12	30
Cerebellum, fusiform gyrus	R	159	4.80	34	-65	-16
			4.62	29	-68	-21
			3.71	27	-67	-32

Alas, no brain area showed a specific sensitivity to pattern violations in this multiple regression RSA with six trial types. As noted above, this could be due to the fact that, when the modality or the cue changed, no area registered the fact that the abstract pattern also changed. To test this idea, we conducted a secondary multiple regression RSA that included only the control standard and the pure pattern deviant trials, with DSMs for modality, cue, pattern and pure pattern violation. Several regions were now sensitive to a pure pattern violation (Figure 7B, Table 5). Most significant effects were in bilateral IPS, bilateral IFG/preCG, and bilateral middle/superior temporal gyrus. Smaller but still significant effects were also present in right MOG, left aINS, left medial frontal gyrus, SMA, anterior MFG, right precuneus, left SFG and right posterior MTG. These results are thus quite consistent with the univariate analyses showing that unexpected changes in the serial order of the sequence elements are reflected in IFG, IPS, SMA and STG regions.

Table 5: Significant activation clusters for the pure pattern violation DSM in the secondary multiple regression RSA searchlight analysis with deviant trials (voxel-wise $p < .001$ unc. and cluster-wise $p < .05$, FDR-corr)

Region	H	k	t	x	y	z
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), middle occipital gyrus, supramarginal gyrus	L	4702	8.49	-38	-42	44
			7.51	-33	-56	45
			6.64	-55	-42	26
Middle temporal gyrus, superior temporal gyrus	R	283	8.00	45	-23	-7
			5.41	48	-35	-9
Inferior frontal gyrus (pars opercularis), precentral gyrus, middle frontal gyrus	L	1979	7.97	-50	16	40
			5.77	-54	14	23
			5.52	-45	21	37
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	R	3098	7.37	50	11	35
			6.99	52	23	7
			6.63	43	14	33
Middle temporal gyrus	L	317	6.86	-52	-44	3
			4.45	-57	-23	-7
			4.23	-61	-37	3
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), supramarginal gyrus, angular gyrus	R	2546	6.68	24	-70	54
			6.34	38	-56	45
			6.10	41	-51	40
Middle occipital gyrus, superior occipital gyrus	R	340	6.37	34	-75	26
			5.75	31	-81	33
			4.98	29	-72	17
Insula, inferior frontal gyrus (pars triangularis, pars orbitalis)	L	282	6.32	-33	28	5
			4.60	-29	16	-2
Precuneus	R	112	5.73	8	-68	54
Superior frontal gyrus (medial)	L	142	5.58	-3	35	47
			4.72	-3	26	40
Supplementary motor area	L/R	951	5.57	8	14	56
			5.28	1	18	66
			5.13	6	9	72
Middle frontal gyrus	L	95	5.32	-38	40	14
Superior frontal gyrus	L	149	5.13	-22	-2	66
			4.80	-10	4	80
			3.66	-20	-5	58
Middle frontal gyrus, superior frontal gyrus	R	475	5.10	36	47	16
			4.75	18	63	24
			4.64	39	51	28
Middle temporal gyrus	R	117	4.40	60	-54	5
			4.14	46	-53	0

Conversely, a final RSA analysis focused on the trials with modality and cue violations, and examined if it made a difference whether they had an additional pattern violation. We included DSM regressors for modality, cue and pattern, pattern violation and violation type (modality deviants or cue deviants). As in the corresponding univariate analysis, no area showed a significant sensitivity to the presence of abstract pattern violations.

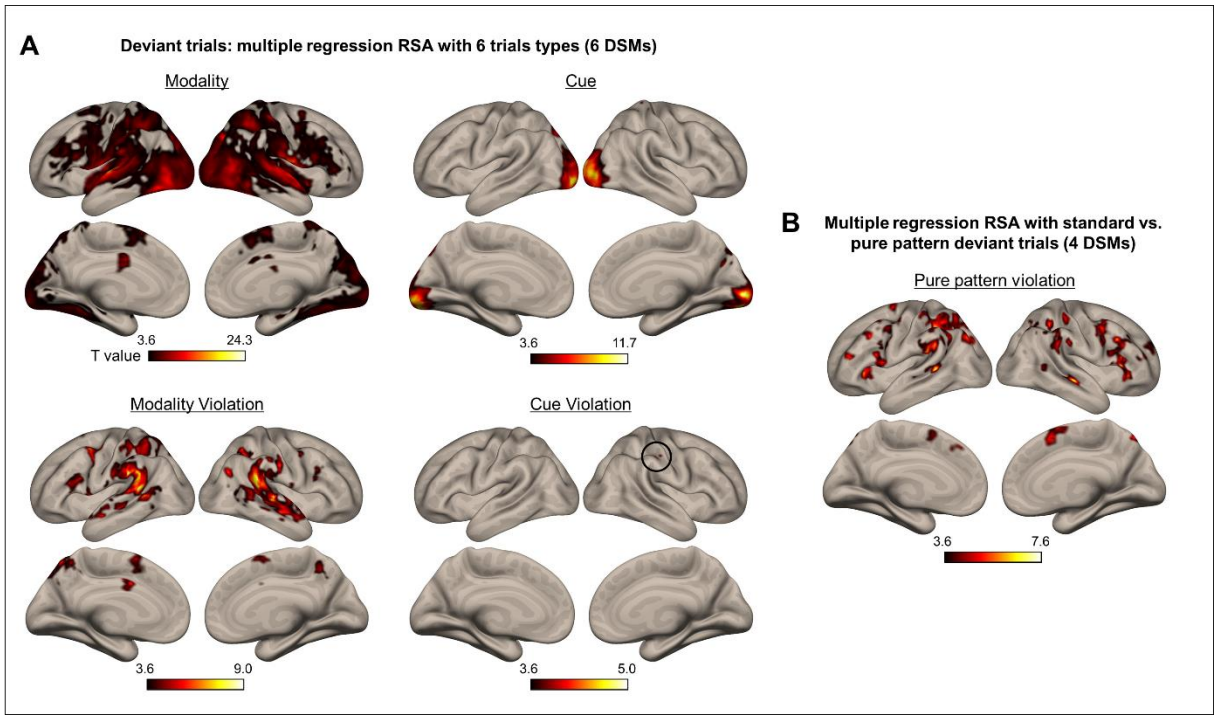


Figure 7: Representational similarity (RSA) analyses of deviant trials. The empirical similarity matrices in a searchlight sphere were modeled by a multiple regression with 6 dissimilarity matrices (DSM). **A**) Significant effects observed with 4 of the 6 DSMs in the main analysis including all 6 trial types. **B**) Significant effects for the "pure pattern violation" DSM" in the secondary analysis including 2 trial types (standard and pure pattern deviants). Threshold for all maps: voxel-wise p < .001 unc. and cluster-wise p < .05 FDR corr.

3.2.4. ROI analyses

Pattern violation effects

In order to test whether language-related areas responded to a pattern change, we assessed the significance of the main pure pattern violation effect (i.e. data extracted from the contrast "pattern deviant > control standard", across all conditions) in subject-specific language-responsive voxels within each of 7 language-related ROI using one-sample t-tests. Since the same procedure was applied for the two other networks, a Hochberg-Bonferroni correction (for 25 ROIs) was applied to the p values (Hochberg, 1988). A positive significant effect was found only for the left IFGoper ($t(18) = 4.27$, p.corr < .006) (Figure 8D, top left). To further test whether this pattern violation effect varied with the modality and the cue of the deviant trial, four contrasts were also examined, corresponding to pure pattern violation effect with visual-spatial, visual-identity, auditory-spatial and auditory-identity trials). For each ROI, the contrast estimates for each participant in each of the four contrasts were entered in a linear mixed model, with participants as the random factor, and trial modality (two levels) and trial cue (two levels) as fixed factors. ANOVAs performed on each model revealed no main effect of the modality, the cue, nor of the interaction, for none of the 7 ROIs (see Table S5, and figure S3A for pattern

violation effects within each condition), suggesting that these factors were not strong determinants of the amplitude of the effect.

In summary, among the voxels that were activated during sentence processing, it was the left IFG voxels (pars opercularis) that also responded the most clearly to a pure pattern violation in our non-linguistic task.

We performed a similar analysis within subject-specific math-responsive voxels within a math-related set of 7 ROIs. The pure pattern violation effect was significant in all seven ROIs (main effect in Figure 8D, top middle; see figure S3B for pattern violation effects within each condition): right preCG/IFG ($t(18) = 6.59$, $p.corr < .0001$), SMA ($t(18) = 6.45$, $p.corr < .0001$), right IPS ($t(18) = 6.14$, $p.corr < .0001$), left IPS ($t(18) = 5.15$, $p.corr < .002$), left preCG/IFG ($t(18) = 4.21$, $p.corr < .006$), left SFG ($t(18) = 3.91$, $p.corr < .01$) and right SFG ($t(18) = 3.39$, $p.corr < .03$). ANOVAs (see Table S5) revealed a main effect of modality (always visual > auditory) in left MFG ($F(1, 54) = 10.7$, $p < .002$), right MFG ($F(1, 54) = 7.47$, $p < .009$), right IPS ($F(1, 54) = 7.53$, $p < .009$), right SFG ($F(1, 54) = 4.42$, $p < .05$) and SMA ($F(1, 54) = 4.62$, $p < .04$). A significant interaction between cue and modality was also present in SMA ($F(1, 54) = 4.44$, $p < .04$; post-hoc pairwise contrasts only revealed a stronger response in visual-identity as opposed to auditory-identity, $p < .02$).

In summary, in contrast to the language network, all areas involved in mental calculation showed a massive increased activation in response to unexpected pattern changes, suggesting their involvement in distinguishing the temporal patterns AABB (2 groups of two) and ABAB (4 single items). A congruence between the mathematics-related regions and those playing a role in tracking temporal sequences had already been reported by Wang et al. (2019) using geometrical visuospatial sequences. The present data confirms a preference for the visual modality, although sequence violation responses were also observed for auditory sequences and non-spatial sequences, at least for the SMA.

Finally, for completeness, we examine all 11 ROIs of the sequence habituation network for pattern violation effect. The main effect was always significant (corrected $p < .05$) (Figure 8D, top right); indicating that voxels that showed a decreased activation during habituation recovered a strong activation to pure pattern changes: right pSTG/MTG ($t(18) = 7.65$, $p.corr < .0001$), right aINS ($t(18) = 6.09$, $p.corr < .0002$), right IFG/preCG ($t(18) = 5.95$, $p.corr < .0003$), left aINS ($t(18) = 5.83$, $p.corr < .0003$), SMA ($t(18) = 5.65$, $p.corr < .0004$), right aIPS ($t(18) = 5.53$, $p.corr < .0005$), right IFG/MFG ($t(18) = 5.09$, $p.corr < .002$), right pMFG ($t(18) = 4.79$, $p.corr < .003$), left preCG ($t(18) = 4.53$, $p.corr < .004$), left IPS ($t(18) = 4.35$, $p.corr < .005$), right pIPS ($t(18) = 3.49$, $p.corr < .03$). ANOVAs (see Table S5) revealed that 5 ROIs responded more strongly to pattern deviants in the visual modality [main effect of modality in left preCG ($F(1, 54) = 11.6$, $p < .002$), right aIPS ($F(1, 54) = 6.46$, $p < .02$), right pIPS ($F(1,$

54) = 5.04, p < 0.03), right IFG/preCG (F(1, 54) = 4.13, p < .05) and SMA (F(1, 54) = 4.20, p < 0.05)]. One ROI, the right pSTG/MTG, was more sensitive to auditory deviants (F(1, 54) = 4.23, p < .05). Only the right pMFG and left and right aINS did not show any effect of modality or cue (p > .10).

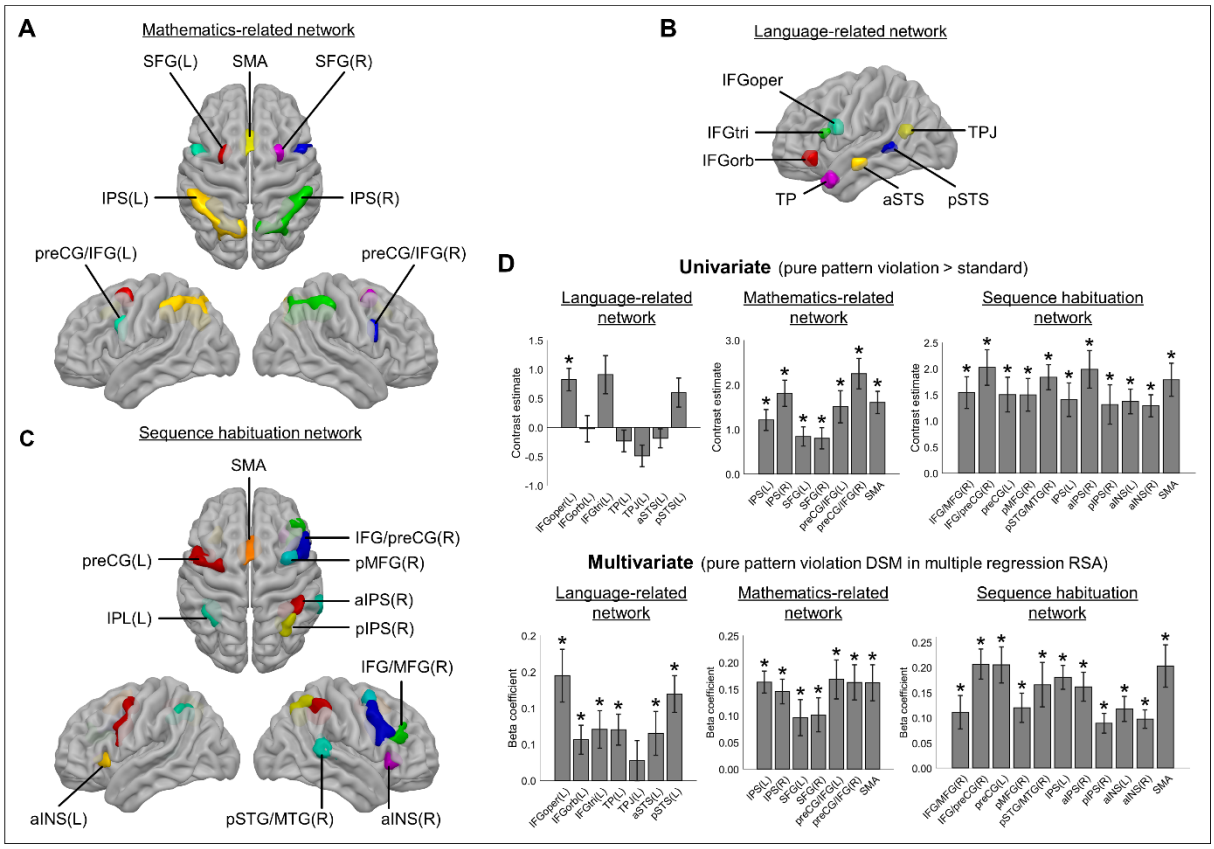


Figure 8: Responses to pure pattern violations in several independently and individually defined regions of interest (ROIs). **A)** Mathematics-related network (7 ROIs). **B)** Language-related network (7 ROIs). **C)** Sequence habituation networks. **D)** Pure pattern violation effects in the three networks of ROIs: "pattern deviant > control standard" univariate contrast (main effect across conditions) (top) and beta coefficients for the "pure pattern violation" DSM in the multivariate, multiple regression RSA (that also included DSMs for modality, cue and pattern) performed over 16 trial types (8 control standards and 8 pure pattern deviants) (bottom). Black star: corrected p < .05, gray star: uncorrected p < .05.

Effects of the type of violation

At the ROI level, we again asked if some regions uniquely responded to pattern changes, or if they also responded to pure modality violations or to pure cue violations. For each ROI, data from 12 contrasts (3 violation types × 4 trial types) were included in a mixed model, with factors of modality, cue and violation type (including participants as the random factor). It is important to keep in mind that it is the modality and cue of the deviant trial that are considered here, not the modality or cue of the block in which the violation occurred. Only the main effects involving the violation type are commented here (all effects are summarized in Table S6) and are plotted in Figure 9 (Figure S4 and Figure S5 show them within each condition).

In the language-related network, all 7 regions responded equally to all three types of violations, and no significant main effect of violation type effect was observed. Note that IFGoper, which showed a strong pure pattern violation effect, also showed similar pure modality ($t(18) = 4.94$, $p.corr < .004$) and pure cue violation effects ($t(18) = 5.12$, $p.corr < .003$) (Figure 9A), with no significant difference between violation types. Similarly, most ROIs of the mathematics-related network showed pattern, modality and cue violation effects (Figure 9B). ANOVAs including the three types of violations in each condition revealed that the amplitude of the response significantly varied depending of the violation type for only two of them: left IPS ($F(2, 200) = 6.37$, $p < .003$; lower response to pattern violations than to modality or cue violations) and SMA ($F(2, 200) = 6.41$, $p < .003$; higher response to modality violations than to pattern or cue violations). Finally, similar observations were made in the ROIs of the habituation network, where most violations yielded a significant response (Figure 9C). The amplitude of the violation effects did vary with the type of violation for five of them: right pSTG/MTG ($F(2, 200) = 14.96$, $p < .0001$; stronger response to modality violations and reduced response to pattern violations), left IPS ($F(2, 200) = 6.12$, $p < .003$; reduced response to pattern violations), right pIPS ($F(2, 200) = 3.34$, $p < .04$; reduced response to pattern violations), left aINS ($F(2, 200) = 4.67$, $p < .02$; reduced response to cue violations) and SMA ($F(2, 200) = 3.47$, $p < .04$; stronger response for modality violations). Interactions of the type of violation with modality were also present, reflecting stronger modality violation effects with visual deviant trials: in right pMFG ($F(2, 200) = 5.04$, $p < .008$), right pIPS ($F(2, 200) = 6.12$, $p < .003$) and left IPS ($F(2, 200) = 4.24$, $p < .02$).

In summary, whether it was in the language-network, the mathematics network or the habituation network, novelty responses were never selective to pure pattern violation, but responded to any unexpected event, compatible with a role in encoding or maintaining a detailed representation of the standard sequence and identifying any of its violations.

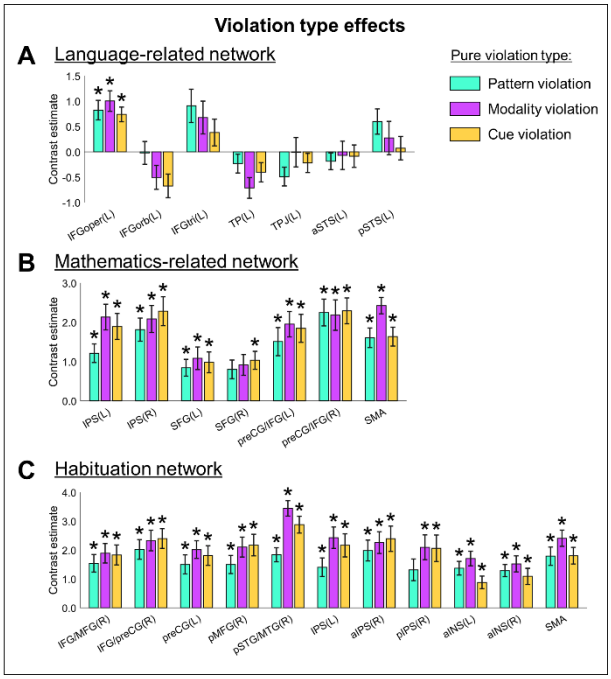


Figure 9: Amplitude of each pure violation effect in various ROIs. (**A**) Language-related network; (**B**) mathematics-related network; (**C**) sequence habituation network (**C**). *: corrected p < .05 (Hochberg-Bonferroni correction for 25 ROIs × 3 pure violation types).

We also reassessed, within each ROIs and for each trial type, the existence of an additional response when the pattern was violated, on top of a modality or cue violation (see Figures S6 and S7). No significant effect was found in T-tests after applying the Hochberg-Bonferroni correction procedure for 25 ROIs × 4 conditions.

Finally, we also reassessed all of these issues using ROI-based multivariate analyses (identical to the searchlight analyses reported above). Detailed results are presented in Appendix, Table S7, and Figure 8D. In summary, we again failed to find any ROI reliably encoding the abstract pattern of sequences or maintaining a common representation when all but the pattern changed. However, again several areas (especially the pSTS language area, IPS, preCG/IFG, SMA and aINS) were highly sensitive to a change in sequence pattern.

4. Discussion

An implicit learning of abstract algebraic patterns is evident in language, mathematics or music, and has been shown to exist in infants (Marcus et al., 1999) and even non-human primates (Shima et al., 2007; Wang et al., 2015). However, experiments typically only a single modality of stimulus presentation (e.g. auditory syllables). Here, by measuring fMRI responses to pattern violations across four sequence presentation formats using a 2 x 2 design with factors of modality, visual or auditory, and cue, spatial or non-spatial, we aimed to clarify whether the adult human brain contains regions that code for abstract, modality- and cue-independent sequence patterns such as AABB, or whether there are only concrete modality- or cue-specific cerebral codes. We also aimed to examine the relation of those responses to the brain regions that encode linguistic and mathematical structures. Here we discuss those two goals in turn.

4.1. Abstract, domain-general versus specific representations of sequences

In the introduction, we proposed 5 ideal criteria for a region coding for abstract algebraic patterns: (1) habituating when the pattern is repeated; (2) dishabituating when the pattern is violated; (3) doing so regardless of the modality or cue in which the pattern is presented, and (4) independently of where other changes also co-occur; and (5) containing decodable information about the patterns. In reality, only the first 3 criteria converged. We now summarize and discuss those complex findings.

Concerning habituation, when presenting the same sequence of stimuli repeatedly over time, we indeed observed a signal decrease, not only in modality-specific areas (visual when the sequences were visual and auditory when they were auditory), but also in a network with a very similar activation profile across conditions: right IFG and right IPS, but also SMA, aINS and cerebellum. This finding is consistent with past results suggesting a crucial role for IFG in the encoding of linguistic (e.g. Bahlmann et al., 2008; Dehaene et al., 2015; Friederici, Fiebach, et al., 2006; Petersson et al., 2012) as well as non-linguistic sequence patterns (e.g. Maess et al., 2001; Wang et al., 2015, 2019). Could this domain-general sequence habituation network indicate a decreasing involvement of the attentional system, which involves very similar regions (Corbetta & Shulman, 2002; Eckert et al., 2009; Fox et al., 2006)? Although a non-specific attentional component cannot be excluded, there are some inconsistencies with this single interpretation. For instance, the IPS has been described as primarily concerned with top-down, goal-directed attention, in the dorsal attentional network, although no task was asked of our participants. It is also systematically described as involved in conjunction with the frontal eye fields, in the dorsal premotor cortex, which was absent here. The IFG or ventral frontal cortex is on the opposite considered as part of the ventral attentional network usually devoted to the detection of

salient stimuli, but in association with the temporoparietal junction, also absent from our network. The involvement of the cerebellum here is furthermore indicative of some sensory prediction or sequence detection activity (Molinari et al., 2008; Nixon, 2003).

Concerning response to violations, our paradigm included pure violations of the temporal pattern governing the sequence (analogous to the global violation effect in the local-global paradigm; Bekinschtein et al., 2009; Wacongne et al., 2011). Behaviorally, although no active detection of deviants was requested from our participants, the pupillometry results (with the data obtained on a subsample of 11 participants) indicate that these violations were detected and triggered a novelty response (in accordance with similar results reported in the auditory modality by Quirins et al., 2018). The corresponding brain response, consistent with habituation results, involved inferior frontal, superior temporal, premotor and intraparietal regions (Figure 5A). However, we found substantial differences in the recruited network depending on the presentation format. Pure pattern violations in the auditory modality were characterized by clear bilateral MTG/STG activation. With the exception of a small posterior region in the visual-spatial condition, probably related to the processing of motion (Pelphrey et al., 2005; Warren & Griffiths, 2003), there was no such temporal activation in the visual modality, which included a right ITG activation.

In many other areas, however, the pattern violation effects were independent of modality and cue (Figure 5B). Contrary to the findings of Michalka et al. (2015), we did not find a clear segregation between areas predominantly devoted to auditory processing and to visual or spatial processing. Instead, a bilateral network for pattern violations was identified. It involved prefrontal, intraparietal, middle and superior temporal areas, as well as the SMA and the anterior insula. All areas of the habituation network also showed a strong response to the main contrast for pattern violations. Furthermore, three areas were found in the conjunction of four independent contrasts for pattern violation within each of the cues and modalities: right IFG, SMA, and right STG. Finally, a multiple regression RSA, accounting for the modality and cue factors, revealed that the pure pattern violation effect concerned a large set of regions, that included bilateral IPS, bilateral IFG/preCG and bilateral STG/MTG.

It thus appears that, by our criteria 1, 2 and 3, the right IFG/inferior premotor region is the main region showing both a habituation effect and a consistent response to pattern violations across all conditions. Our data therefore is consistent with the assumption of a crucial role of this region in abstract representation of non-linguistic rules (Badre et al., 2010; Wang et al., 2015, 2019; Wilson et al., 2015). Concerning its right-hemisphere lateralization, although several authors argue for a role of Broca's area in the left hemisphere for the processing of hierarchical temporal structures (Bahlmann

et al., 2008; Friederici, Bahlmann, et al., 2006; Koechlin & Jubault, 2006; Makuuchi et al., 2009), the right IFG is also often found active, particularly when using non-linguistic spatial or music-like stimuli (Maess et al., 2001; Wang et al., 2015, 2019). The IPS, which appeared in the above contrasts whenever spatial sequences were involved (Figure 5B), is also a likely contributor to this abstract representation of sequences. Past research found it responsive to number changes (Wang et al., 2015), with an activity correlates with the complexity of visual sequences (Wang et al., 2019), and it is occasionally reported as part of the network responding to rule violations (Bekinschtein et al., 2009; Schröger et al., 2007; Wacongne et al., 2011). The present data suggest that it might play a more systematic role, independent of the modality or sequence presentation formats.

By introducing additional pure modality and pure cue violations trials, where the pattern was maintained while the modality or the cue changed, we could question the specificity of these regions to pure pattern violations. Overall, highly similar regions of the bilateral parietal, prefrontal and temporal cortex responded to pattern, cue and modality violations (Figure 5). Thus, these areas form a global novelty response network that responds to any mismatch between the predicted stimulus and the observed one, regardless of the type of mismatch. This finding is in accordance with the hypothesis that a high-level, amodal global workspace network reacts to consciously perceived novelty (Bekinschtein et al., 2009; Dehaene et al., 1998; Dehaene & Changeux, 2011; Strauss et al., 2015). The suggestion is that this network encodes a summary representation of all aspects of conscious experience, thus including both abstract pattern as well as sensory details. We nevertheless identified some regions (dorsal part of the preCG/SFG and bilateral pars orbitalis of the IFG) that responded more strongly to pure pattern violations (Figure 6). Pars orbitalis of the IFG, especially in the left hemisphere, is usually considered as primarily concerned with high-level semantic and syntactic aspects of language (Pallier et al., 2011; Price, 2012; Santi & Grodzinsky, 2010). Superior frontal regions (close to the SMA) have also been linked to syntactic violations (Newman et al., 2001). The present result may thus indicate that these areas may have a more fundamental function in encoding the abstract temporal pattern of sequences, rather than in responding to any unexpected stimuli. This might be related with the assumption that the more anterior frontal regions support increasingly abstract representations (Badre et al., 2010; Badre & D'Esposito, 2007; Koechlin & Jubault, 2006).

Criterion 4, our most demanding test for a shared representation of abstract patterns across presentation formats, asked whether a brain region reacted whenever the pattern changed, even if the modality or cue also changed. However, this criterion was never met, preventing us to definitely conclude that a purely abstract and generalizable representation of the pattern was encoded and used for prediction. Instead, it appears that, whenever the modality or the cue changed, the ensuing novelty response was so large as to swamp any other putative additional novelty due to pattern change.

Behaviorally, the increase in pupil diameter following the onset of a modality or cue deviant trial did not significantly vary whether the abstract pattern was the same as the standard or not (although it should be kept in mind that these analyses were only conducted on a subsample of 11 participants). Regarding brain activation, only limited effects were found in the anterior cingulate gyrus when an auditory-spatial trial occurred in a visual-spatial block and in the SMA/paracentral lobule when a visual-identity trial occurred in a visual-spatial block. However, no generic effect was found, and this null effect was also the outcome of RSA analyses. Although it is difficult to give an interpretation for an absence of effect, the fact that behavioral, univariate and multivariate analyses converged suggests that the premise of our criterion 4 could have been wrong. Even if abstract, modality- and cue independent pattern representations exist, they may not generate an additional signal on top of the huge evidence for violation which arises when either the cue or the modality changes. According to a prominent theory (Botvinick & Watanabe, 2007; McCoy et al., 2019; Smolensky, 1990), the neural representation of sequences could result from the binding, through a tensor-product operation, of two distinct neural vectors, one for the abstract structures (here the patterns AABB or ABAB) and another for the filler item (here the specific images or sounds that fill the slots A and B and exemplify the pattern). If this was the case, then a change in any of these vectors would suffice to completely change the neural vector, and there would be no additional novelty response when the pattern changes on top of other changes, exactly as observed here.

Finally, our fifth criterion asked whether any region allowed to decode which of the two abstract patterns (AABB and ABAB) was presented: did they show stronger responses to one pattern or the other in a univariate analysis, or a pattern-based similarity in a multivariate analysis? Unfortunately, the latter analysis, based on RSA searchlight, did not reveal any area specifically sensitive to the sequence pattern itself, within either standard trials or deviant trials. This null finding is inconclusive and could simply reflect the fact that different sequence patterns are encoded by tightly intermingled neural populations in IFG or SMA (Fujii & Graybiel, 2003; Shima et al., 2007; Shima & Tanji, 2006) which cannot be discriminated at the current resolution of fMRI. In the univariate analysis, only a ventral precuneus area was found. Activation was reduced in this region for the alternating (ABAB) relative the two-pair pattern (AABB). Thus, it may be related to the different demands that those two patterns made on the tracking of item number, frequencies and/or transition probabilities, since this area was reported to be part of the network involved in the learning of such sequence statistics (Giorgio et al., 2018). However, since this area was not part of either the habituation or the pure pattern violation network, it did not meet our criteria 1-3, and hence its role in pattern coding remains unclear.

4.2. Relation to language- and math-related networks

Gary Marcus (Marcus, 2019; Marcus et al., 1999) has emphasized how the encoding of algebraic patterns may be an important step in a hierarchy of increasingly abstract sequence processing abilities, culminating in high-level human skills such as language and mathematics (see also Dehaene et al., 2015). For instance, experiments using artificial grammars have shown that Broca's area, which is involved in the processing of language-like sequential structures (Bahlmann et al., 2008; Fitch & Friederici, 2012; Friederici, Bahlmann, et al., 2006), is also involved in the encoding of short binary auditory patterns (Wang et al., 2015). Furthermore, overlapping areas of dorsal prefrontal and parietal cortices are activated when performing simple mathematical operations and when processing visuospatial sequences (Wang et al., 2019). Here, using subject-specific ROI analyses with functionally and individually-defined language and mathematics networks, we thus hoped to improve our understanding of the relationships between sequence processing and these high-level functions, in particular by characterizing what relates to the presentation format of sequence (auditory, visual, spatial or identity-based), and what relates more to the encoding of abstract structure.

We found that all the areas of the math-related network, activated when subjects made simple arithmetic calculations, were also sensitive to pure pattern violations (Figure 8). The effect was stronger in the visual modality for most of them, particularly for visuo-spatial sequences (Figure S2B), which could be expected considering they already showed a sensitivity to this modality within standard trials (Figure 3). As previously described, they also responded to other types of violations (Figure 9), indicating that they encoded more than just the sequence's algebraic pattern, or at least responded to novelty in a non-specific way. Mathematics-related areas, in the IPS or the prefrontal cortex, are also thought to support representations of space, quantity, time and temporal order (Amalric & Dehaene, 2016; Harvey et al., 2015; Nieder, 2012), which are all dimensions that can be involved in the encoding of our sequences. Thus, the present results suggest that, although these regions may be particularly apt at encoding the visual, numerical and spatial dimensions of sequences, these capacities are deployed even for auditory sequences.

By contrast, only a few areas of language-related network consistently showed a clear response to pure pattern violations (Figure 8). Indeed, although the effect reached significance in the highly sensitive RSA for most of them (except TPJ), only the left IFGoper showed a significant activation increase in univariate analyses correcting for the number of analyzed ROIs. This finding confirms the central role of the frontal inferior language-responsive cortex in encoding sequence structure, and its capacity to do, within the very same subject-specific voxels, even for non-linguistic stimuli (as advocated earlier, Bahlmann et al., 2008; Huettel et al., 2002; Koechlin & Jubault, 2006; Wang et al.,

2015). Interpretation should be cautious, since this region also strongly responded to pure modality- and cue-violation events (Figure 9A), in accordance with the idea that it is sensitive to all sorts of regularities (Dehaene et al., 2015; Huettel et al., 2002).

Overall, our results suggest that, although both language and math networks were recruited by our non-linguistic sequence paradigm, they did so to a different extent: Mathematics-related areas were more globally engaged than language-related ones in detecting our sequence changes. This may be due to their sensitivity to space, time or number, which are thought to be critical features in the abstract language needed to account for human non-linguistic sequence encoding in working memory (Planton et al., 2021). Nevertheless, the inferior frontal region (especially pars opercularis), with a right lateralization, previous known to play a crucial role in syntactic processing, is reinforced by the present study in its role for extracting and/or storing pattern information from sequences of stimuli. Future research will be required to examine the respective roles of these regions, but the present findings are compatible with the general hypothesis that IFG may "bind" or "unify", in a structured syntactic frame, the stimulus-specific representations encoded in more posterior brain regions (Hagoort, 2013).

4.3. Limitations and future extensions

The present work suffers from several limitations, which also point to ways in which the search for abstract sequence patterns in the human brain could be improved in future experiments. Here we briefly mention three of the most salient ones.

First, we only tested two extremely short and simple sequence patterns: AABB and ABAB. It seems likely that such sequences do not require much effort to be encoded in working memory. In a recent study (Planton et al., 2021), we found that humans use a complex language with a nested, recursive structure ("repetitions of repetitions") in order to encode binary sequences and compress them in memory. However, we also found that they do so primarily when the sequences are long and complex, such that they can be memorized only after they are compressed (e.g. 16-item sequences such as AAAABBBBAABBABAB). Shorter sequences, as used here, may simply be committed to memory without any sophisticated compression or structural encoding. In future work, the present design should be replicated with longer sequences of variable compressibility. We would predict that clearer fMRI evidence for abstract patterns should emerge for long yet regular and therefore compressible sequences.

Second, our goal here was to test whether a *passive* presentation paradigm could uncover the cerebral correlates of abstract sequence encoding in human subjects. Clearly, an active task, such as sequence recall, could have been better, inasmuch as it would have forced subjects to achieve a deeper

level of abstract encoding. The present work was influenced by the proposal that humans have a propensity to automatically infer hierarchical structures, especially nested tree structures, from any temporal sequence (i.e. the dendrophilia hypothesis; Fitch, 2014). However, the systematic, spontaneous, automatic and implicit nature of this ability remains to be proven, and may turn out to be less prominent than initially thought. Instead, a conscious effort may be needed to achieve a highly abstract level of sequence representation. This is also an important issue for comparative studies aiming at comparing human versus non-human primate sequence processing abilities, or at distinguishing innate and acquired sequence abilities in infants. Animal and infant studies (e.g. Basirat et al., 2014; Gil-da-Costa et al., 2013; Wang et al., 2015; Wilson et al., 2015) often have to rely on passive novelty-response paradigms with a relatively short learning phase, similar to the one used here.

A third limitation is that, in a given block, a *very simple fixed pattern* was presented, without variations in item identity (unlike e.g. Wang et al., 2015). Under such conditions, we did not find any strong evidence that an abstract pattern was encoded in the brain. Memory remained grounded in the multidimensional, but essentially sensory, representation of the sequence, presumably because such a representation was entirely sufficient to predict the future sequence items. Varying the stimuli during the habituation period may provide a better methodology in order to induce the participant to search for a higher-level rule (Wang et al., 2015).

Finally, since this work targeted abstract and amodal effects by using multiple sequence presentations formats (e.g. seeking activations underlying shared representations), the analyses focused on a single presentation format, thus using only a subset of the data, may have lacked statistical power. This may have prevented us from uncovering effects restricted to a given modality or presentation format. Since differences in the way human processes sequences in different modalities have been reported (e.g. Freides, 1974; Patel et al., 2005) affecting for instance the ease of learning, experimental designs focusing on a single format may perhaps be better suited to assess some of the most subtle effects that we have failed to demonstrate here. The sensitivity to such effects should also be improved by focusing on predefined and individually localized regions, as has been done here for the regions of language and mathematics. Yet, a functional localizer paradigm dedicated to the localization of (non-linguistic) sequence processing areas and controlling for various confounds remains to be elaborated and validated.

Overall, our results indicate that, even in humans, implicit learning of an abstract pattern may not occur spontaneously, but may require additional conditions, possibly including longer and more

complex sequences, an active task, and exposure to multiple exemplars of the rule; such a design, could be used in future studies.

Conclusion

The present results, while complex and not definitely conclusive, once again emphasize that the human inferior frontal gyrus as one of the key sites for the encoding of the abstract structure of temporal sequences, even outside the language domain. Novelty responses, which were exploited here to a maximum, are useful but provide only an indirect probe of the internal representation of sequences in this area. In the future, we plan to use more direct methods, including intracranial recordings in human and non-human primates, in order to clarify how neural ensembles coordinate in order to encode abstract sequence representations.

Appendix

Multiple regression RSAs with ROIs

Similar to the searchlight analyses, two series multiple regression RSAs were conducted using voxels within each ROI for each participant. The analyses for standard trials used 16 maps (one for each unique standard trial type) and 3 dissimilarity matrices (DSM) in the regression ("modality", "cue", "pattern"). The analyses for deviant trials used 48 maps (one for each deviant trial type) and 6 dissimilarity matrices (DSM) in the regression ("trial modality", "trial cue", "trial pattern", "modality violation", "cue violation", "pattern violation"). Note that, for the latter, the modality cue and pattern of the deviant trial where considered, not the modality cue and pattern of the block in which they were presented. Significance of the betas weights corresponding to each DSM was assessed using one-sided one-sample t-tests (using Hochberg-Bonferroni correction for 25 ROIs). A summary of significant effects obtained in all multiple regression RSAs is provided in Table S7.

Regarding standard trials, no sensitivity to modality, cue or pattern were found for any area of the left-hemispheric language-related network. In the mathematics-related network, a sensitivity to the modality was present for 4 out of 7 ROIs: left IPS ($t(18) = 4.68$, $p.corr < .003$), right IPS ($t(18) = 3.90$, $p.corr < .02$), right preCG/IFG ($t(18) = 3.72$, $p.corr < .02$) and left preCG/IFG ($t(18) = 3.50$, $p.corr < .03$). Only the left IPS showed a sensitivity to the cue ($t(18) = 3.34$, $p.corr < .05$) and none to the pattern. In the sequence habituation network the beta weights for the modality DSM were significant in 7 ROIs: right IFG/preCG ($t(18) = 3.66$, $p.corr < .02$), left preCG ($t(18) = 5.07$, $p.corr < .001$), right pSTG/MTG ($t(18) = 6.09$, $p.corr < .0001$), left IPS ($t(18) = 3.41$, $p.corr < .03$), right aIPS ($t(18) = 3.29$, $p.corr < .04$) and right pIPS ($t(18) = 3.95$, $p.corr < .02$). No area showed an effect of the cue or the pattern.

Before conducting the main RSA with all deviant trial types, as in the whole brain analysis, in order to focus specifically on the pure pattern violation effect, a secondary multiple regression RSAs was conducted by only including control standard trials and pure pattern deviants (included DSMs were "modality", "cue", "pattern" and "pure pattern violation"). As shown in Figure 8D (bottom left) a sensitivity to pure pattern violations was found for all language-related ROIs but TPJ: especially pSTS ($t(18) = 4.73$, $p.corr < .002$) and IFGoper ($t(18) = 3.97$, $p.corr < .005$), but also TP ($t(18) = 3.30$, $p.corr < .02$), IFGorb ($t(18) = 2.83$, $p.corr < .03$), IFGtri ($t(18) = 2.73$, $p.corr < .03$) and aSTS; ($t(18) = 2.16$, $p.unc < .05$). Although most of these ROIs did not showed differences in average activation amplitude between standard and pure pattern deviants, the multivariate approach thus revealed that they actually encoded information regarding the sequence pattern. As expected, since they all responded to pure pattern violation in the univariate analysis, all ROIs of the mathematics network also showed the effect:

left IPS (t(18) = 8.00, p.corr < .0001), right IPS (t(18) = 6.31, p.corr < .0001), right preCG/IFG (t(18) = 4.97, p.corr < .001), SMA (t(18) = 4.79, p.corr < .002), left preCG/IFG (t(18) = 4.62, p.corr < .002), right SFG (t(18) = 3.20, p.corr < .02) and left SFG (t(18) = 2.84, p.corr < .03) (Figure 8D, bottom middle). It was also the case for each ROI of the habituation network (Figure 8D, bottom right): right IFG/preCG (t(18) = 6.94, p.corr < .0001), left IPS (t(18) = 7.69, p.corr < .0001), right aIPS (t(18) = 5.64, p.corr < .0003), left preCG (t(18) = 5.73, p.corr < .0003), right aINS (t(18) = 5.39, p.corr < .0004), SMA (t(18) = 4.80, p.corr < .002), left aINS (t(18) = 4.72, p.corr < .002), right pIPS (t(18) = 4.56, p.corr < .002), right pMFG (t(18) = 4.14, p.corr < .004), right pSTG/MTG (t(18) = 3.77, p.corr < .007), right IFG/MFG (t(18) = 3.34, p.corr < .02).

In the main multiple regression RSA, including all deviant trial types, all language-related areas showed a sensitivity to the modality of the deviant trial: IFGoper (t(18) = 5.86, p.corr < .0001), pSTS (t(18) = 4.0, p.corr < .004), TPJ (t(18) = 3.61, p.corr < .007), IFGtri (t(18) = 3.37, p.corr < .008), IFGorb (t(18) = 3.32, p < .008), aSTS (t(18) = 3.22, p.corr < .008) and TP (t(18) = 2.79, p.corr < .02). Five were sensitive to the presence of a modality violation: IFGoper (t(18) = 5.60, p.corr < .0004), aSTS (t(18) = 4.70, p.corr < .008), IFGtri (t(18) = 4.15, p.corr < .004), pSTS (t(18) = 3.76, p.corr < .008) and TPJ (t(18) = 3.02, p.corr < .02). None was sensitive to trial cue or pattern. None to a violation of the cue or of the pattern.

A clear sensitivity to the modality of deviant trials was found for each area of the mathematics-related network: left preCG/IFG (t(18) = 9.02, p.corr < .0001), right preCG/IFG (t(18) = 9.01, p.corr < .0001), left IPS (t(18) = 7.29, p.corr < .0001), right IPS (t(18) = 7.68, p.corr < .0001), left SFG (t(18) = 6.15, p.corr < .0001), right SFG (t(18) = 4.02, p.corr < .004) and SMA (t(18) = 5.66, p.corr < .0001). All also encoded the presence of a modality violation: left IPS (t(18) = 5.51, p.corr < .0004), right IPS (t(18) = 4.29, p.corr < .004), left preCG/IFG (t(18) = 4.31, p.corr < .004), right preCG/IFG (t(18) = 4.24, p.corr < .004), SMA (t(18) = 4.30, p.corr < .004), right SFG (t(18) = 2.78, p.corr < .04 and left SFG (t(18) = 2.62, p.corr < .04. Only left IPS encoded the cue (t(18) = 3.65, p.corr < .03). None was sensitive to the specific pattern but left IPS encoded the presence of pattern violations (t(18) = 3.93, p.corr < .02).

In the sequence habituation network, all areas were sensitive to the modality of the deviant trials: right aIPS (t(18) = 8.27, p.corr < .0001), right IFG/preCG (t(18) = 9.13, p.corr < .0001), left IPS (t(18) = 7.50, p.corr < .0001), right pIPS (t(18) = 6.67, p.corr < .0001), left preCG (t(18) = 10.52, p.corr < .0001), right pSTG/MTG (t(18) = 7.74, p.corr < .0001), SMA (t(18) = 7.76, p.corr < .0001), right pMFG (t(18) = 5.04, p.corr < .0006), right IFG/MFG (t(18) = 4.72, p.corr < .001), right aINS (t(18) = 3.51, p.corr < .008) and left aINS (t(18) = 2.45, p.corr < .02). Most areas were sensitive to a violation of the modality: left IFG/preCG (t(18) = 6.22, p.corr < .0001), right pSTG/MTG (t(18) = 7.97, p.corr < .0001), SMA (t(18) =

5.18, p.corr < .0008), left IPS (t(18) = 4.91, p.corr < .002), right IFG/preCG (t(18) = 4.61, p.corr < .003), right IFG/MFG (t(18) = 4.01, p.corr < .006), right pIPS (t(18) = 3.93, p.corr < .006), right aIPS (t(18) = 3.48, p.corr < .02), left aINS (t(18) = 3.12, p.corr < .03) and right pMFG (t(18) = 2.92, p.corr < .03). Only left preCG was sensitive to the cue (t(18) = 3.88, p.corr < .02) and none to cue violations. None was sensitive to the pattern but the pattern violation factor was significant for SMA (t(18) = 4.01, p.corr < .02).

Across the three networks, a few ROIs were found to encode the presence of abstract pattern violations. Since all 6 trial types were included in the analysis, this may be the result of a cross-modal (or cross-cue) generalization of the pattern (with modality and cue deviants), but also to an effect of the simple detection of pure pattern change (with pure pattern deviants). In order to disentangle these two accounts, a second secondary RSA was performed using only modality and cue deviants. Here, no significant effect of pattern violation was found for any ROI. This suggests that when the set of stimuli changed (modality or cue change), the abstract pattern of standard could not be decoded.

Declaration of interest

None

Funding

This research was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM, http://www.inserm.fr), the Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA, http://www.cea.fr), the Collège de France (https://www.college-de-france.fr/site/college/index.htm), the Bettencourt-Schueller Foundation (http://www.fondationbs.org) and a European Research Council (ERC, https://erc.europa.eu/) grant to S.D. ("NeuroSyntax", ID: 695403). This research was also partly supported by the European Union (FEDER-2007–2013, agreement #91–2015–004).

CRediT author statement

Samuel Planton: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization, Project administration. **Stanislas Dehaene**:

Conceptualization, Methodology, Resources, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.

Others

• No part of the study procedures or analyses was pre-registered prior to the research being conducted.

• The conditions of the ethics approval do not permit public archiving of study data. Readers seeking access to the data should contact the first or last author. Access will be granted depending on completion of a formal data sharing agreement, proper measures taken for protection of privacy and approval by the relevant ethical committee.

• We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

• Legal copyright restrictions prevent public archiving of the various instruments used in the current study, which can be obtained from the copyright holders in the cited references

• Stimuli, experiment code and analysis code can be found here: https://osf.io/cw7hx/.

References

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., 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

Babkoff, H., Muchnik, C., Ben-David, N., Furst, M., Even-Zohar, S., & Hildesheimer, M. (2002). Mapping lateralization of click trains in younger and older populations. *Hearing Research*, *165*(1), 117–127. https://doi.org/10.1016/S0378-5955(02)00292-7

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

Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition*, *132*(2), 137–150. https://doi.org/10.1016/j.cognition.2014.03.013

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

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

Botvinick, M., & Watanabe, T. (2007). From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *The Journal of Neuroscience: The Official Journal*

of the Society for Neuroscience, *27*(32), 8636–8642. https://doi.org/10.1523/JNEUROSCI.2110-07.2007

Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M.-J., Tian, B., Rauschecker, J. P., & Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, *2*(8), 759–766. https://doi.org/10.1038/11239

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755

Davis, T., & Poldrack, R. A. (2013). Measuring neural representations with fMRI: Practices and pitfalls. *Annals of the New York Academy of Sciences*, *1296*, 108–134. https://doi.org/10.1111/nyas.12156

Dehaene, S., & Changeux, J.-P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(24), 14529–14534. https://doi.org/10.1073/pnas.95.24.14529

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

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*(3), 277–283. https://doi.org/10.1038/72991

Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2009). At the heart of the ventral attention system: The right anterior insula. *Human Brain Mapping*, *30*(8), 2530–2541. https://doi.org/10.1002/hbm.20688

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

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. (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 of London. Series B, Biological Sciences*, *367*(1598), 1933–1955. https://doi.org/10.1098/rstb.2012.0103

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science (New York, N.Y.)*, *303*(5656), 377–380. https://doi.org/10.1126/science.1089401

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046–10051. https://doi.org/10.1073/pnas.0604187103

Freides, D. (1974). Human information processing and sensory modality: Cross-modal functions, information complexity, memory, and deficit. *Psychological Bulletin*, *81*(5), 284–310. https://doi.org/10.1037/h0036331

Friederici, A. D., Bahlmann, J. r., 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*, *103*(7), 2458–2463. https://doi.org/10.1073/pnas.0509389103

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cereb Cortex*, *16*(12), 1709–1717. https://doi.org/10.1093/cercor/bhj106

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

Gil-da-Costa, R., Stoner, G. R., Fung, R., & Albright, T. D. (2013). Nonhuman primate model of schizophrenia using a noninvasive EEG method. *Proceedings of the National Academy of Sciences*, *110*(38), 15425–15430. https://doi.org/10.1073/pnas.1312264110

Giorgio, J., Karlaftis, V. M., Wang, R., Shen, Y., Tino, P., Welchman, A., & Kourtzi, Z. (2018). Functional brain networks for learning predictive statistics. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *107*, 204–219. https://doi.org/10.1016/j.cortex.2017.08.014

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*, 416. https://doi.org/10.3389/fpsyg.2013.00416

Hartmann, M., & Fischer, M. H. (2014). Pupillometry: The Eyes Shed Fresh Light on the Mind. *Current Biology*, *24*(7), R281–R282. https://doi.org/10.1016/j.cub.2014.02.028

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*(44), 13525–13530. https://doi.org/10.1073/pnas.1515414112

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

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, *75*(4), 800–802. https://doi.org/10.1093/biomet/75.4.800

Huettel, S. A., Mack, P. B., & McCarthy, G. (2002). Perceiving patterns in random series: Dynamic processing of sequence in prefrontal cortex. *Nature Neuroscience*, *5*(5), 485–490. https://doi.org/10.1038/nn841

Kabdebon, C., & Dehaene-Lambertz, G. (2019). Symbolic labeling in 5-month-old human infants. *Proceedings of the National Academy of Sciences*, *116*(12), 5805–5810. https://doi.org/10.1073/pnas.1809144116

Kasper, L., Bollmann, S., Diaconescu, A. O., Hutton, C., Heinzle, J., Iglesias, S., Hauser, T. U., Sebold, M., Manjaly, Z.-M., Pruessmann, K. P., & Stephan, K. E. (2017). The PhysIO Toolbox for Modeling Physiological Noise in fMRI Data. *Journal of Neuroscience Methods*, *276*, 56–72. https://doi.org/10.1016/j.jneumeth.2016.10.019

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

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. https://doi.org/10.1073/pnas.0600244103

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 4. https://doi.org/10.3389/neuro.06.004.2008

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress & L. A. (Ed) Jeffress (Eds.), *Cerebral mechanisms in behavior; the Hixon Symposium.* (1952-04498-003; pp. 112–146). Wiley.

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*, *131*(5), 1332–1343. https://doi.org/10.1093/brain/awn040

Lewis, J. W., Beauchamp, M. S., & DeYoe, E. A. (2000). A Comparison of Visual and Auditory Motion Processing in Human Cerebral Cortex. *Cerebral Cortex*, *10*(9), 873–888. https://doi.org/10.1093/cercor/10.9.873

Linden, D. E., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F. E., Goebel, R., & Dierks, T. (1999). The functional neuroanatomy of target detection: An fMRI study of visual and auditory oddball tasks. *Cerebral Cortex (New York, N.Y.: 1991)*, *9*(8), 815–823. https://doi.org/10.1093/cercor/9.8.815

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nat Neurosci*, *4*(5), 540–545.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences*, *106*(20), 8362–8367. https://doi.org/10.1073/pnas.0810928106

Marcus, G. F. (2019). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. MIT Press.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science (New York, N.Y.)*, *283*(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

McCoy, R. T., Linzen, T., Dunbar, E., & Smolensky, P. (2019). RNNs Implicitly Implement Tensor Product Representations. *ArXiv:1812.08718 [Cs]*. http://arxiv.org/abs/1812.08718

Michalka, S. W., Kong, L., Rosen, M. L., Shinn-Cunningham, B. G., & Somers, D. C. (2015). Short-Term Memory for Space and Time Flexibly Recruit Complementary Sensory-Biased Frontal Lobe Attention Networks. *Neuron*, *87*(4), 882–892. https://doi.org/10.1016/j.neuron.2015.07.028

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

Näätänen, R. (2003). Mismatch negativity: Clinical research and possible applications. *International Journal of Psychophysiology*, *48*(2), 179–188. https://doi.org/10.1016/S0167-8760(03)00053-9

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An Event-Related fMRI Study of Syntactic and Semantic Violations. *Journal of Psycholinguistic Research*, *30*(3), 339–364. https://doi.org/10.1023/A:1010499119393

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proceedings of the National Academy of Sciences*, *109*(29), 11860–11865. https://doi.org/10.1073/pnas.1204580109

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

Nourski, K. V., Steinschneider, M., Rhone, A. E., Kawasaki, H., Howard, M. A., & Banks, M. I. (2018). Processing of auditory novelty across the cortical hierarchy: An intracranial electrophysiology study. *NeuroImage*, *183*, 412–424. https://doi.org/10.1016/j.neuroimage.2018.08.027

Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-Modal Multivariate Pattern Analysis of Neuroimaging Data in Matlab/GNU Octave. *Frontiers in Neuroinformatics*, *10*, 27. https://doi.org/10.3389/fninf.2016.00027

Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *88*(2), 109–123. https://doi.org/10.1016/j.ijpsycho.2013.03.015

Paavilainen, P., Jaramillo, M., Näätänen, R., & Winkler, I. (1999). Neuronal populations in the human brain extracting invariant relationships from acoustic variance. *Neuroscience Letters*, *265*(3), 179–182. https://doi.org/10.1016/S0304-3940(99)00237-2

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, *108*(6), 2522–2527. https://doi.org/10.1073/pnas.1018711108

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, *163*(2), 226–238. https://doi.org/10.1007/s00221-004-2159-8

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional Anatomy of Biological Motion Perception in Posterior Temporal Cortex: An fMRI Study of Eye, Mouth and Hand Movements. *Cerebral Cortex*, *15*(12), 1866–1876. https://doi.org/10.1093/cercor/bhi064

Petersson, K.-M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*(2), 83–95. https://doi.org/10.1016/j.bandl.2010.08.003

Pinel, P., Thirion, B., Meriaux, S., Jobert, A., Serres, J., Le Bihan, D., Poline, J.-B., & Dehaene, S. (2007). Fast reproducible identification and large-scale databasing of individual functional cognitive networks. *BMC Neuroscience*, *8*, 91. https://doi.org/10.1186/1471-2202-8-91

Planton, S., Kerkoerle, T. van, 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

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Quirins, M., Marois, C., Valente, M., Seassau, M., Weiss, N., El Karoui, I., Hochmann, J.-R., & Naccache, L. (2018). Conscious processing of auditory regularities induces a pupil dilation. *Scientific Reports*, *8*(1), 14819. https://doi.org/10.1038/s41598-018-33202-7

Saarinen, J., Paavilainen, P., Schöger, E., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport*, *3*(12), 1149–1151. https://doi.org/10.1097/00001756-199212000-00030

Santi, A., & Grodzinsky, Y. (2010). FMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage*, *51*(4), 1285–1293. https://doi.org/10.1016/j.neuroimage.2010.03.034

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

Shima, K., & Tanji, J. (2006). Binary-coded monitoring of a behavioral sequence by cells in the pre-supplementary motor area. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *26*(9), 2579–2582. https://doi.org/10.1523/JNEUROSCI.4161-05.2006

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, *46*(1), 159–216. https://doi.org/10.1016/0004-3702(90)90007-M

Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., Naccache, L., van Wassenhove, V., & Dehaene, S. (2015). Disruption of hierarchical predictive coding during sleep. *Proceedings of the National Academy of Sciences*, *112*(11), E1353–E1362. https://doi.org/10.1073/pnas.1501026112

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior Formation of Cortical Memory Traces for Melodic Patterns in Musicians. *Learning & Memory*, *8*(5), 295–300. https://doi.org/10.1101/lm.39501

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

Wacongne, C., Labyt, E., Wassenhove, V. van, Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *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. (2019). 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

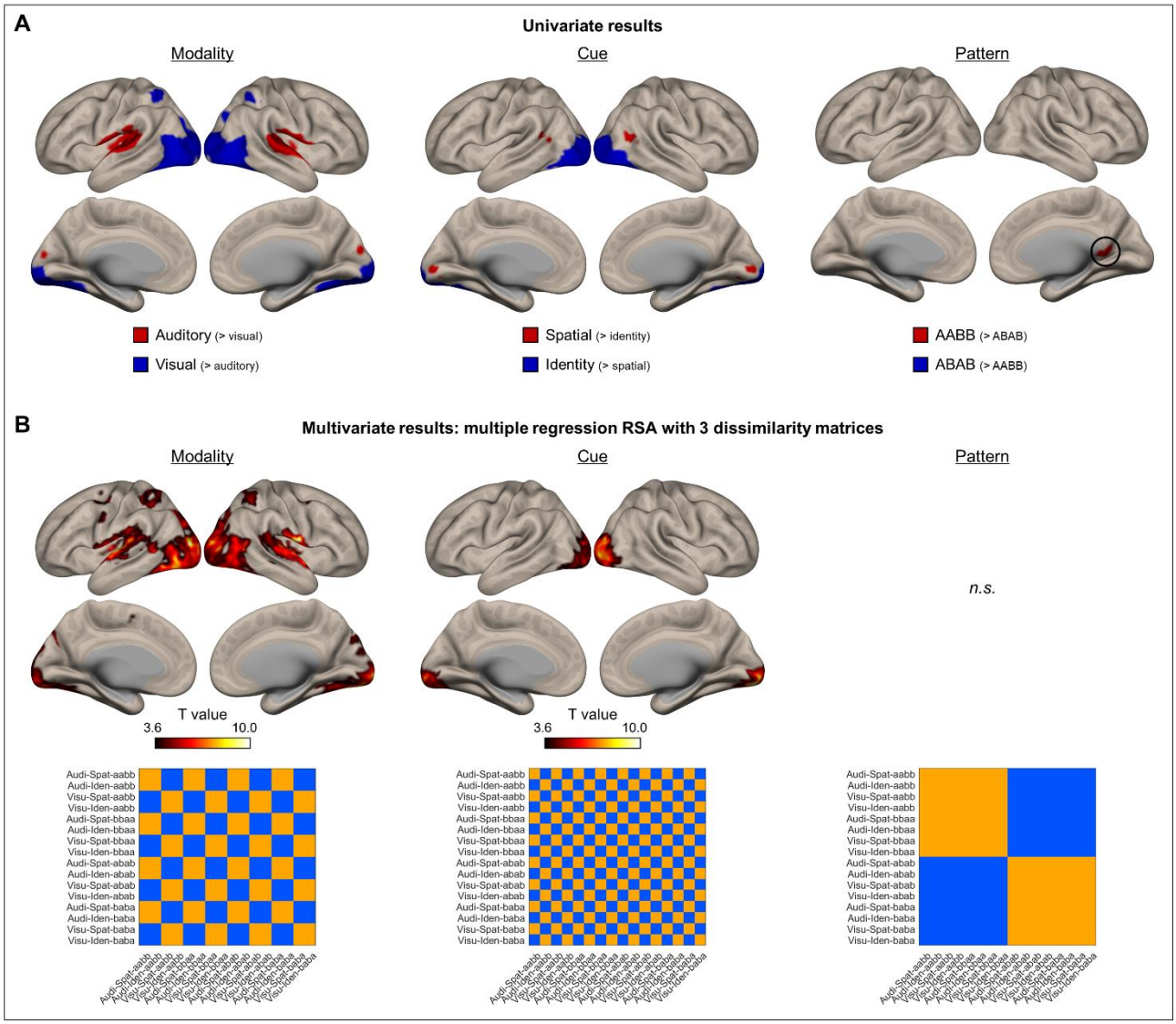
Warren, J. D., & Griffiths, T. D. (2003). Distinct Mechanisms for Processing Spatial Sequences and Pitch Sequences in the Human Auditory Brain. *Journal of Neuroscience*, *23*(13), 5799–5804. https://doi.org/10.1523/JNEUROSCI.23-13-05799.2003

Warren, J. D., Zielinski, B. A., Green, G. G. R., Rauschecker, J. P., & Griffiths, T. D. (2002). Perception of Sound-Source Motion by the Human Brain. *Neuron*, *34*(1), 139–148. https://doi.org/10.1016/S0896-6273(02)00637-2

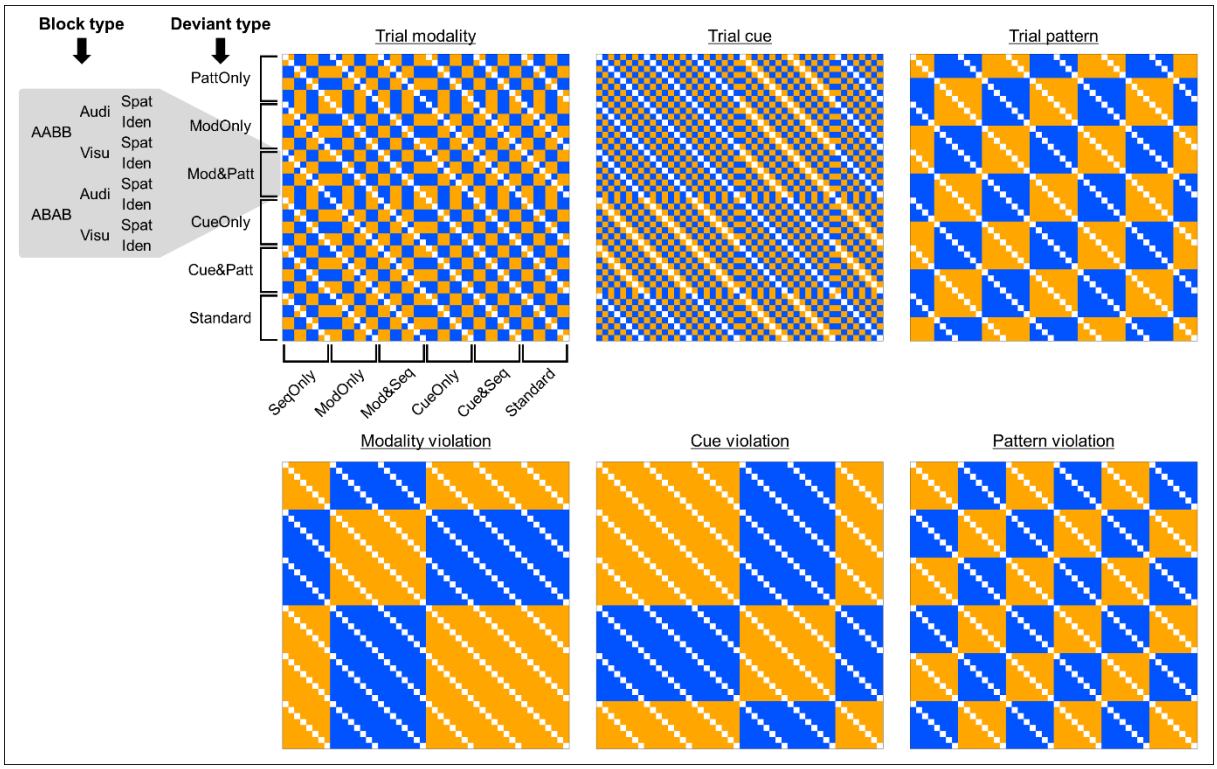
Wilson, B., Kikuchi, Y., Sun, L., Hunter, D., Dick, F., Smith, K., Thiele, A., Griffiths, T. D., Marslen-Wilson, W. D., & Petkov, C. I. (2015). Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature Communications*, *6*(1). https://doi.org/10.1038/ncomms9901

Zylberberg, A., Oliva, M., & Sigman, M. (2012). Pupil dilation: A fingerprint of temporal selection during the “attentional blink.” *Frontiers in Psychology*, *3*, 316. https://doi.org/10.3389/fpsyg.2012.00316

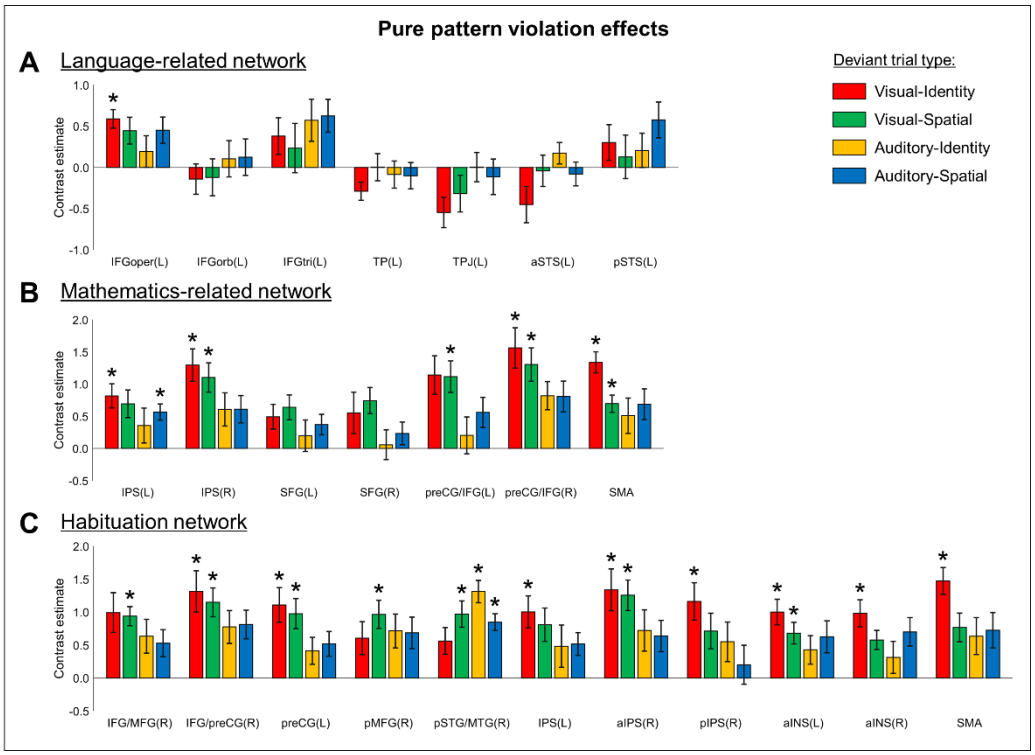
5. Supplementary Materials



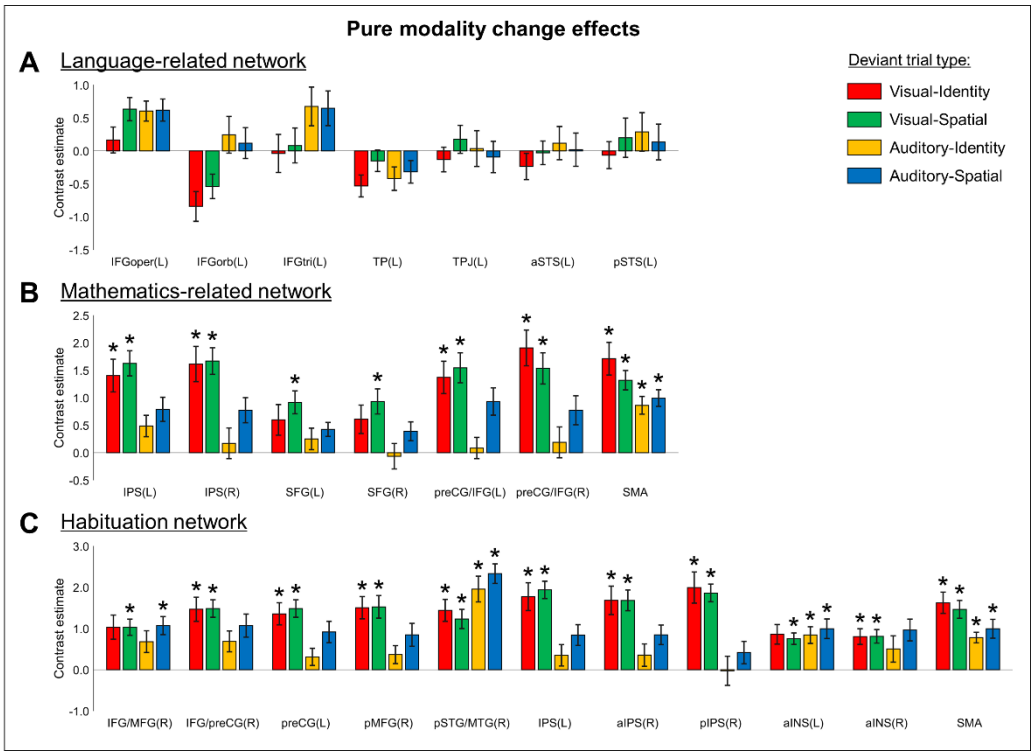
5.1. Figure S10: Main conditions effects in standard trials. A) Univariate contrasts for the effets of modality (left), cue (middle) and pattern (right). B) Results of the multivariate multiple linear regression RSA for the modality (left), cue (middle) and pattern DSMs. Corresponding DSMs are shown in the bottom (blue represents similarity, orange represents dissimilarity). All activation maps are reported with voxel-wise ($p < .001$, uncorrected) and cluster-wise ($p < .05$, FDR corrected) thresholds.



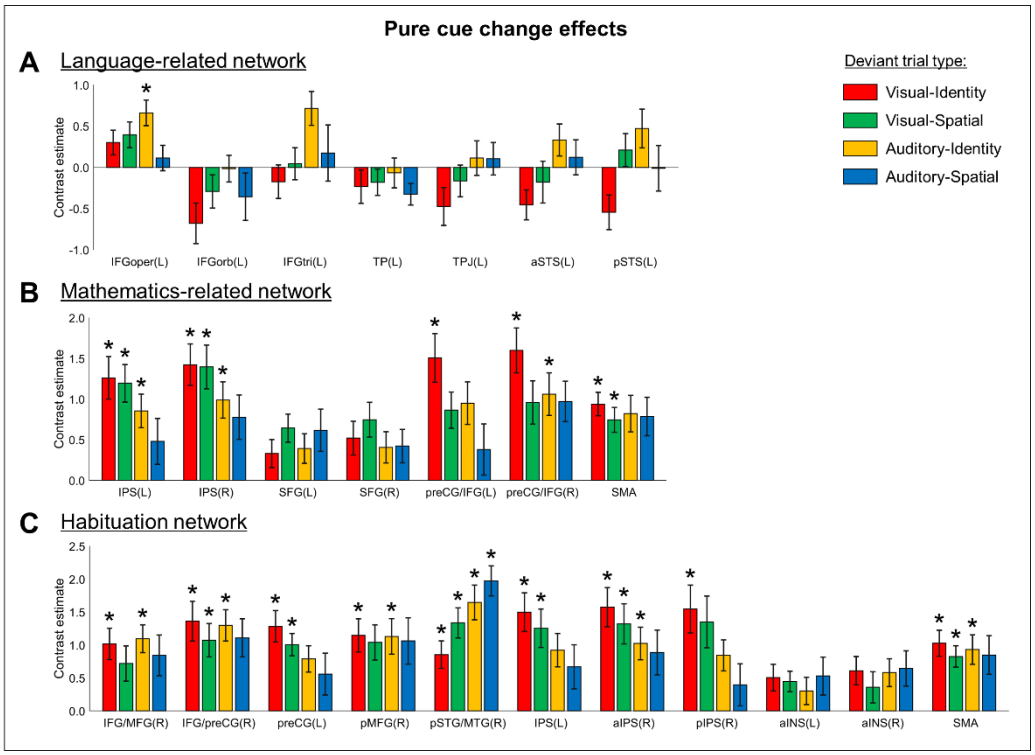
5.2. Figure S11: Dissimilarity matrices used in the multiple regression RSA with deviant trials. Each of the 5 deviant types (as well as 1 control standard) occurred in each of the 8 different block types (e.g. Audi-Spat-AABB, Audi-Iden-AABB, etc.), resulting in 48×48 matrices. Note that the "Trial modality', "Trial Cue" and "Trial pattern" DSMs refer to the modality/cue/pattern of the deviant trial, not of the block type in which they were presented. Blue represents similarity, orange represents dissimilarity. The white cells show the comparisons excluded from the analyses, as they correspond to the correlations between trials from the same blocks. PattOnly: pure pattern violation; ModOnly: pure modality violation; Mod&Patt: modality and pattern violation; CueOnly: pure cue violation; Cue&Patt: cue and pattern violation.



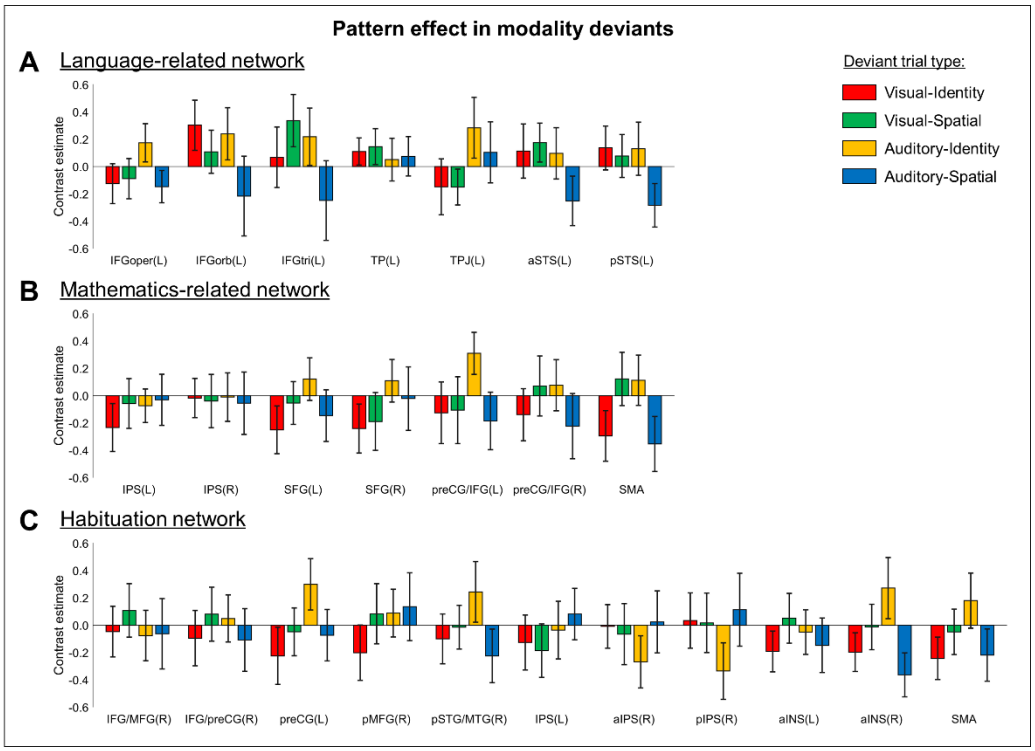
5.3. Figure S12: Pure pattern change effects. Amplitude of the pure pattern violation effect ("pattern deviant > control standard" contrast) in the three sets of individually and functionally-defined ROIs, depending on the deviant trial type. *: Hochberg-Bonferroni-corrected $p < .05$. Error bars represent SEM.



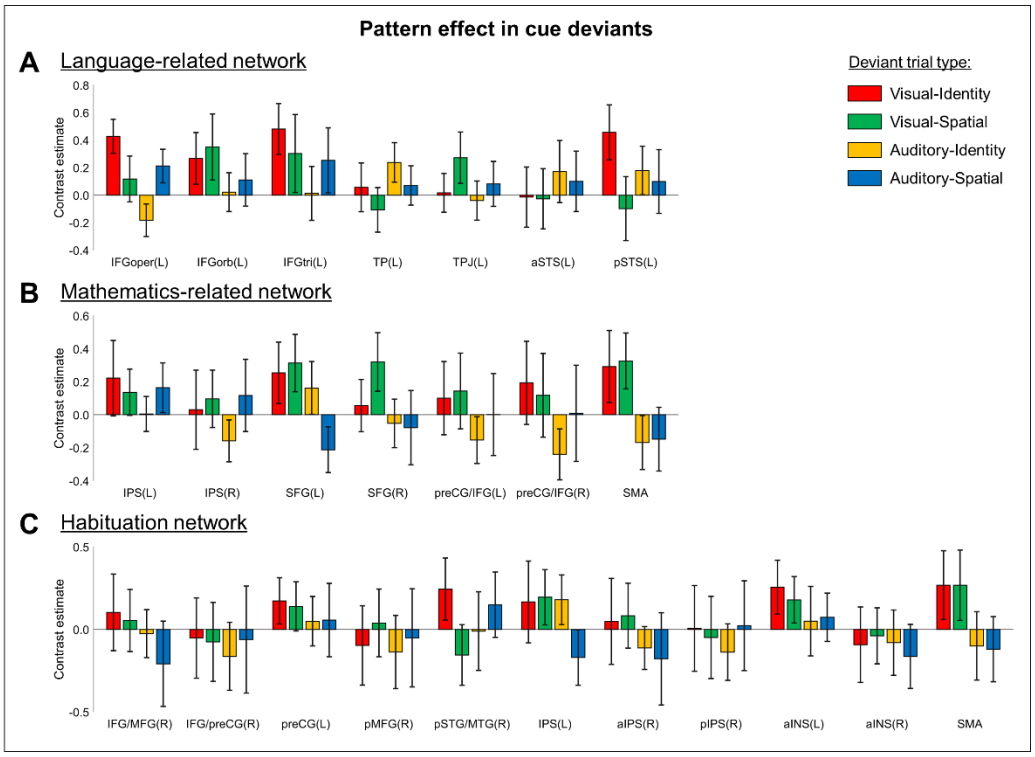
5.4. Figure S13: Pure modality change effects. Amplitude of the pure modality violation effect ("pure modality deviant > same-block control standard" contrast) in the three sets of individually and functionally-defined ROIs, depending on the deviant trial type*: Hochberg-Bonferroni-corrected $p < .05$. Error bars represent SEM.



5.5. Figure S14: Pure cue change effects. Amplitude of the pure cue violation effect ("pure cue deviant > same-block control standard" contrast) in the three sets of individually and functionally-defined ROIs, depending on the deviant trial type. *: Hochberg-Bonferroni-corrected $p < .05$. Error bars represent SEM.



5.6. Figure S15: Pattern effect in modality deviants. Amplitude of the pattern effect in modality deviants (i.e. "modality+pattern deviant > pure modality deviant")in the three sets of individually and functionally-defined ROIs, depending on the deviant trial type. No significant effect was found after applying Hochberg-Bonferroni correction. Error bars represent SEM.



5.7. Figure S16: Pattern effect in cue deviants. Amplitude of the pattern effect in cue deviants (i.e. "cue+pattern deviant > pure cue deviant")in the three sets of individually and functionally-defined ROIs, depending on the deviant trial type. No significant effect was found after applying Hochberg-Bonferroni correction. Error bars represent SEM.

5.8. Table S1: Significant activation clusters for the "activation decrease" contrast in habituation in each of the four trial-type conditions (voxel-wise p<.001 unc. and cluster-wise p < .05, FDR-corr)

Region	H	k	t	x	y	z
Activation decrease in habituation: Auditory-spatial condition						
Superior temporal gyrus, middle temporal gyrus, anterior insula, supramarginal gyrus	R	6626	7.62	64	-35	10
			6.86	46	-40	9
			6.80	55	-30	14
Superior temporal gyrus, middle temporal gyrus, supramarginal gyrus	L	3735	7.27	-45	-26	7
			6.95	-45	-39	21
			6.11	-61	-32	19
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus		3645	6.18	50	7	49
			5.78	48	16	26
			5.74	43	7	59
Anterior insula	L	619	5.51	-31	26	2
			4.51	-31	25	12
Supplementary motor area	L/R	1357	5.30	-5	4	59
			5.00	4	16	49
			4.52	-8	-5	70
Precentral gyrus	L	130	4.85	-55	4	33
Precentral gyrus, middle frontal gyrus	L	335	4.64	-38	-2	51
			3.79	-52	0	51
			3.18	-40	-2	61
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), supramarginal gyrus	R	647	4.59	36	-46	49
			4.29	43	-37	44
			3.52	45	-51	61
Inferior parietal gyrus	L	183	4.55	-34	-42	40
Posterior cerebellum	L	702	4.51	-10	-79	-46
			4.34	-31	-63	-27
			4.31	-17	-74	-25
Caudate	R	294	4.42	13	7	10
			4.33	24	2	9
			3.77	6	-4	9
Activation decrease in habituation: Auditory-identity condition						
Superior temporal gyrus, middle temporal gyrus	R	2672	6.70	67	-32	3
			5.80	55	-18	5
			4.98	55	11	-13
Superior temporal gyrus, middle temporal gyrus	L	784	5.70	-48	-23	5
			4.77	-36	-33	10
			4.54	-62	-21	0
Inferior frontal gyrus (pars opercularis, pars triangularis), middle frontal gyrus	R	981	5.34	46	12	30
			4.25	41	33	30
			4.19	39	23	21
Posterior cerebellum	L	724	5.29	-13	-74	-39
			4.61	-10	-84	-44
Posterior cerebellum	L	441	4.98	-34	-67	-30
			4.67	-27	-58	-32
Posterior cerebellum	L	170	4.52	-34	-65	-48
			3.74	-41	-77	-42
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), supramarginal gyrus, angular gyrus	R	924	4.44	46	-40	51
			4.06	39	-56	58
			4.03	45	-42	42
Precuneus	L/R	172	4.08	4	-70	54

			3.78	4	-68	66
Supplementary motor area	L/R	127	3.73	4	14	56
			3.42	8	26	44
Inferior frontal gyrus (pars triangularis, pars orbitalis), anterior insula	R	163	3.68	36	32	0
			3.48	43	25	-2
Activation decrease in habituation: Visual-spatial condition						
Fusiform gyrus, lingual gyrus, inferior occipital gyrus, middle occipital gyrus, cerebellum	L	2120	6.34	-12	-91	-2
			5.75	-34	-67	-21
			5.11	-34	-75	-18
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus	R	2563	6.33	43	4	58
			4.89	43	4	31
			4.54	46	16	31
Fusiform gyrus, lingual gyrus, inferior occipital gyrus, middle occipital gyrus, cerebellum	R	1674	6.08	15	-89	-2
			5.41	29	-82	-14
			4.77	36	-68	-16
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), angular gyrus, precuneus	R	3547	5.88	29	-63	45
			5.59	25	-53	44
			5.18	15	-67	44
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus	L	1183	5.74	-31	-5	54
			4.76	-52	2	31
			4.43	-38	9	24
Supplementary motor area	L/R	766	5.52	-8	11	52
			5.08	6	12	54
			4.78	-8	2	58
Anterior insula	L	336	5.16	-27	23	2
			5.14	-27	32	3
Anterior insula	R	513	4.76	32	25	5
			3.88	36	18	-2
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), precuneus	L	1631	4.71	-33	-46	38
			4.64	-22	-60	35
			4.55	-31	-56	51
Middle frontal gyrus	R	218	4.64	39	46	37
			4.08	43	39	38
			3.56	36	33	31
Middle occipital gyrus	L	238	4.27	-24	-82	17
Cingulate gyrus	L/R	170	4.15	6	-33	26
			3.85	-5	-32	26
Activation decrease in habituation: Visual-identity condition						
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	L	3618	8.02	-40	2	31
			5.47	-27	0	58
			5.10	-45	28	28
Fusiform gyrus, inferior occipital gyrus, middle occipital gyrus, cerebellum, inferior temporal gyrus, lingual gyrus	L/R	14672	7.61	-34	-63	-14
			7.23	36	-72	-13
			7.22	32	-65	-21
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus)	L	3734	7.26	-31	-56	51
			5.86	-38	-44	47
			5.43	-40	-39	40
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	R	3083	7.11	48	11	35
			5.46	39	18	24
			5.34	36	2	59
	R	3331	7.03	31	-58	52

Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus), angular gyrus, superior occipital gyrus			6.46	29	-70	30
			6.43	31	-49	51
Supplementary motor area	L/R	1503	6.29	-5	12	51
			6.23	-6	5	59
			4.59	8	11	56
Anterior insula	L	564	5.64	-31	19	2
Cerebellum	R	570	4.93	29	-70	-53
			4.70	36	-61	-48
			4.28	22	-75	-49
Anterior insula	R	476	4.77	31	21	9
			4.39	38	21	-4
Middle frontal gyrus	L	587	4.69	-40	56	17
			4.24	-36	47	21
			4.17	-29	49	5
Pallidum, putamen, caudate	L	154	4.52	-19	2	5
			3.44	-15	11	9
Middle/superior frontal gyrus orbital part	L	135	3.92	-20	53	-16
			3.78	-24	44	-21
Caudate, putamen	R	160	3.68	17	11	-2
			3.67	17	0	16
			3.56	15	14	7

5.9. Table S2: Significant activation clusters for the main conditions effects with standard trials (univariate) (voxel-wise p<.001 unc. and cluster-wise p < .05, FDR-corr)

Region	H	k	t	x	y	z
Standard trials (univariate): Auditory > Visual						
Superior temporal gyrus, rolandic operculum	L	4022	15.97	-52	-16	5
Superior temporal gyrus, rolandic operculum	R	4145	14.77	60	-9	5
			13.27	59	-19	7
			12.40	50	-21	9
Cuneus	L/R	148	3.63	6	-86	17
			3.60	-3	-88	16
Standard trials (univariate): Visual > Auditory						
Middle occipital gyrus, inferior occipital gyrus, fusiform gyrus, lingual gyrus	R	8390	11.24	31	-81	-14
			10.86	24	-86	-4
			9.76	31	-86	2
Middle occipital gyrus, inferior occipital gyrus, fusiform gyrus, lingual gyrus	L	8710	10.10	-24	-91	2
			9.92	-41	-72	2
			9.69	-34	-79	-16
Intraparietal sulcus	L	411	5.56	-26	-53	52
Middle occipital gyrus, inferior occipital gyrus	R	337	5.47	27	-77	31
Intraparietal sulcus	R	354	4.87	27	-49	52
			3.97	31	-63	56
Standard trials (univariate): Spatial > Identity						
Lingual gyrus, calcarine gyrus	L	189	5.79	-12	-88	-2
Lingual gyrus, calcarine gyrus	R	149	5.02	11	-88	0
Middle temporal gyrus	R	224	4.79	43	-61	12
			4.02	55	-63	9
Middle temporal gyrus	L	365	4.78	-57	-63	14
			3.81	-54	-53	14
Lingual gyrus, cerebellum	R	111	4.45	10	-75	-13
			3.17	3	-77	-7
Standard trials (univariate): Identity > Spatial						
Middle occipital gyrus, inferior occipital gyrus, fusiform gyrus, lingual gyrus	R	4074	11.81	32	-88	2
			8.77	20	-98	2
			6.73	39	-72	-9
Middle occipital gyrus, inferior occipital gyrus, fusiform gyrus, lingual gyrus	L	3904	10.51	-36	-86	-11
			9.84	-27	-88	-4
			8.17	-20	-100	-4
Standard trials (univariate): AABB > ABAB						
Precuneus, posterior cingulate gyrus, lingual gyrus	R	209	4.19	11	-54	12
			3.63	10	-44	9

5.10. Table S3: Significant activation clusters for the cue and modality DSMs effects in the multiple-regression RSA with standard trials (multivariate) (voxel-wise p<.001 unc. and cluster-wise p < .05, FDR-corr)

Region	H	k	t	x	y	z
Standard trials (multivariate): Modality DSM						
Middle occipital gyrus, inferior occipital gyrus, lingual gyrus, superior temporal gyrus, middle temporal gyrus, fusiform gyrus, cuneus, rolandic operculum, cerebellum	L/R	28963	14.16	-50	-16	3
			11.39	-29	-89	9
			9.97	31	-86	-9
Superior temporal gyrus, rolandic operculum, Heschl's gyrus, temporal pole	R	5281	11.09	62	-7	9
			9.04	59	-23	10
			8.09	50	-11	2
Precentral gyrus, middle frontal gyrus	L	755	7.36	-41	-4	56
			5.86	-31	-5	51
			5.48	-34	-11	44
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus)	L	1023	7.27	-31	-46	47
			5.21	-33	-47	65
			4.84	-27	-60	61
Precentral gyrus, middle frontal gyrus	R	223	5.99	38	0	56
Intraparietal sulcus (inferior parietal gyrus, superior parietal gyrus)	R	817	5.38	31	-53	52
			5.03	18	-63	66
			4.75	25	-61	59
Precentral gyrus, postcentral gyrus	L	150	4.72	-55	-4	33
			4.23	-59	-9	40
			3.90	-59	5	35
Standard trials (multivariate): Cue DSM						
Middle occipital gyrus, inferior occipital gyrus, lingual gyrus, calcarine gyrus	L/R	8584	10.76	13	-91	-7
			10.16	22	-84	-11
			9.10	32	-93	9

5.11. Table S4: Significant activation clusters for the cue and modality violation effects (univariate) (voxel-wise p<.001 unc. and cluster-wise p < .05, FDR-corr)

Region	H	k	t	x	y	z
Violation effects (univariate): modality deviant > (control) standard						
Superior temporal gyrus, middle temporal gyrus, temporal pole, supramarginal gyrus	R	5118	11.34	59	-42	19
			10.82	59	-39	10
			7.56	46	-23	-7
Middle temporal gyrus, superior temporal gyrus, inferior parietal gyrus, superior parietal gyrus, precuneus, angular gyrus, supramarginal gyrus	L/R	12659	10.72	-61	-47	14
			7.93	-29	-53	49
			7.79	32	-54	45
Supplementary motor area	L/R	1731	9.57	-3	11	54
			3.98	-1	26	38
			3.75	-12	4	72
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, insula	L	3508	8.31	-40	0	51
			7.03	-41	2	37
			6.97	-29	28	3
Inferior frontal gyrus (pars opercularis, pars triangularis), middle frontal gyrus, precentral gyrus	R	3676	7.25	43	14	28
			6.93	45	5	44
			6.72	38	4	37
Cerebelum, inferior temporal gyrus	R	1142	5.80	38	-63	-27
			5.12	52	-51	-13
			4.38	43	-47	-14
Posterior cingulate gyrus	L/R	709	5.76	6	-21	28
			5.60	-5	-32	26
			4.96	4	-33	26
Insula	R	457	5.66	34	23	0
Cerebellum	L	471	5.45	-36	-61	-32
			4.86	-34	-68	-27
Cerebellum	R	102	4.93	10	-74	-25
Cerebellum	L	221	4.80	-10	-77	-42
Fusiform gyrus, infgerior temporal gyrus, inferior occipital gyrus	L	502	4.72	-41	-60	-13
			3.89	-34	-47	-20
			3.54	-33	-58	-16
Cerebellum	L	167	4.27	-6	-77	-23
			3.68	-19	-70	-27
Cerebellum	L	118	4.05	-33	-67	-56
Violation effects (univariate): cue deviant > (control) standard						
Superior temporal gyrus, middle temporal gyrus, temporal pole	R	5270	8.84	57	7	-9
			7.79	62	-18	-4
			7.73	66	-33	9
Superior temporal gyrus, middle temporal gyrus, temporal pole	L	4048	8.47	-50	0	-6
			7.58	-64	-21	0
			7.14	-66	-28	5
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus, middle frontal gyrus	R	1414	5.54	46	18	26
			4.23	50	5	45
			4.00	43	39	19
Cerebellum	L	502	5.52	-33	-67	-28
			4.66	-12	-72	-25
			4.11	-13	-77	-41
Angular gyrus, middle occipital gyrus, superior occipital gyrus	R	357	4.80	39	-70	40
			4.42	36	-79	44

Supplementary motor area	L/R	443	4.75	3	4	66
			4.69	-3	12	56
			3.52	6	16	52
Inferior parietal gyrus	L	319	4.69	-38	-44	40
Middle frontal gyrus	R	292	4.63	34	7	58
			3.72	31	11	49
Inferior frontal gyrus (pars opercularis, pars triangularis), precentral gyrus	L	640	4.58	-45	0	23
			4.41	-45	19	19
			4.20	-50	16	26
Inferior parietal gyrus	R	389	4.45	34	-46	38
			3.92	46	-35	51
Precentral gyrus	L	108	3.81	-47	4	49

5.12. Table S5: Modulation of pure pattern violation effects depending on the properties of the main conditions. Significant effects in ANOVAs of with Modality and Cue factors, for each ROI of the three sets of ROIs. Only effects with $p < .10$ are shown. *: $p < .05$, **: $p < .01$.

Language-related network

	Modality	Cue	Interaction
IFGoper (L)	-	-	-
IFGorb (L)	-	-	-
IFGtri (L)	-	-	-
TP (L)	-	-	-
TPJ (L)	-	-	-
aSTS (L)	-	-	-
pSTS (L)	-	-	-

Mathematics-related network

	Modality	Cue	Interaction
IPS (L)	-	-	-
IPS (R)	**	-	-
SFG (L)	-	-	-
SFG (R)	*	-	-
preCG/IFG (L)	**	-	-
preCG/IFG (R)	**	-	-
SMA	*	-	*

Habituation network

	Modality	Cue	Interaction
IFG/MFG (R)	-	-	-
IFG/preCG (R)	*	-	-
preCG (L)	**	-	-
pMFG (R)	-	-	-
pSTG/MTG (R)	*	-	**
IPS (L)	-	-	-
aIPS (R)	*	-	-
pIPS (R)	*	-	-
aINS (L)	-	-	-
aINS (R)	-	-	-
SMA	*	-	-

5.13. Table S6: Modulation of violation effects depending the violation type (i.e., pattern, modality, or cue violation) and the properties of the main condition. Significant effects in ANOVAs of the violation effects, with Violation type, Modality and Cue factors, for each ROI of the three sets of ROIs. Only effects with $p < .10$ are shown. *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Language-related network

	Modality	Cue	Violation type	Modality × Cue	Modality × Viol. type	Cue × Viol. type
IFGoper (L)	-	-	-	-	-	-
IFGorb (L)	***	-	-	-	-	-
IFGtri (L)	***	-	-	-	-	-
TP (L)	-	-	-	-	-	-
TPJ (L)	*	-	-	-	-	-
aSTS (L)	**	-	-	*	-	-
pSTS (L)	-	-	-	-	-	-

Mathematics-related network

	Modality	Cue	Violation type	Modality × Cue	Modality × Viol. type	Cue × Viol. type
IPS (L)	***	-	**	-	-	-
IPS (R)	***	-	-	-	-	-
SFG (L)	*	*	-	-	-	-
SFG (R)	***	-	-	-	-	-
preCG/IFG (L)	***	-	-	-	-	**
preCG/IFG (R)	***	-	-	*	**	-
SMA	**	-	**	*	-	-

Habituation network

	Modality	Cue	Violation type	Modality × Cue	Modality × Viol. type	Cue × Viol. type
IFG/MFG (R)	-	-	-	-	-	-
IFG/preCG (R)	**	-	-	-	-	-
preCG (L)	***	-	-	-	-	-
pMFG (R)	*	-	-	-	**	-
pSTG/MTG (R)	***	-	***	-	-	-
IPS (L)	***	-	**	-	*	-
aIPS (R)	***	-	-	-	-	-
pIPS (R)	***	-	*	-	**	-
aINS (L)	-	-	*	-	-	-
aINS (R)	-	-	-	*	-	-
SMA	***	-	*	-	-	-

5.14. Table S7: Results of ROI-based multiple regression RSAs. Significant effects in one-sample t-tests of beta weights, in the multiple regression RSAs with standard trials and in the three multiple regression RSAs with deviant trials: main (6 trial types), 1st secondary (standard and pure pattern deviants), 2nd secondary (modality and cue deviants with and without pattern violation), i.e., four distinct analyses for each ROI. ***: corrected p < .001, **: corrected p < .01, *: corrected p < .05.

Language-related network

	Standard trials RSA			Main deviant trials RSA						1st secondary RSA[1]	2nd secondary RSA[1]	
	Modality	Cue	Pattern	Modality	Cue	Pattern	Modality viol.	Cue viol.	Pattern viol.	Pure pattern viol.	Violation type	Pattern viol.
IFGoper(L)	-	-	-	***	-	-	***	-	-	**	**	-
IFGorb(L)	-	-	-	**	-	-	-	-	-	*	-	-
IFGtri(L)	-	-	-	**	-	-	**	-	-	*	-	-
TP(L)	-	-	-	*	-	-	-	-	-	*	-	-
TPJ(L)	-	-	-	**	-	-	*	-	-	-	**	-
aSTS(L)	-	-	-	**	-	-	**	-	-	-	*	-
pSTS(L)	-	-	-	**	-	-	**	-	-	**	*	-

Mathematics-related network

	Standard trials RSA			Main deviant trials RSA						1st secondary RSA[1]	2nd secondary RSA[1]	
	Modality	Cue	Pattern	Modality	Cue	Pattern	Modality viol.	Cue viol.	Pattern viol.	Pure pattern viol.	Violation type	Pattern viol.
IPS (L)	**	*	-	***	*	-	**	-	-	***	**	-
IPS (R)	*	-	-	***	-	-	**	-	-	***	*	-
SFG (L)	-	-	-	***	-	-	*	-	-	*	*	-
SFG (R)	-	-	-	**	-	-	*	-	-	*	*	-
preCG/IFG (L)	*	-	-	***	-	-	**	-	-	**	*	-
preCG/IFG (R)	*	-	-	***	-	-	**	-	*	***	*	-
SMA	-	-	-	***	-	-	**	-	-	**	-	-

Sequence habituation network

	Standard trials RSA			Main deviant trials RSA						1st secondary RSA[1]	2nd secondary RSA[1]	
	Modality	Cue	Pattern	Modality	Cue	Pattern	Modality viol.	Cue viol.	Pattern viol.	Pure pattern viol.	Violation type	Pattern viol.
IFG/MFG (R)	-	-	-	***	-	-	**	-	-	*	-	-
IFG/preCG (R)	*	-	-	***	-	-	**	-	-	***	-	-
preCG (L)	***	-	-	***	*	-	***	-	-	***	**	-
pMFG (R)	-	-	-	***	-	-	*	-	-	**	-	-
pSTG/MTG (R)	***	-	-	***	-	-	***	-	-	**	***	-
IPS (L)	*	-	-	***	-	-	**	-	-	***	*	-
aIPS (R)	*	-	-	***	-	-	*	-	-	***	-	-
pIPS (R)	*	-	-	***	-	-	**	-	-	**	-	-
aINS (L)	-	-	-	*	-	-	*	-	-	**	*	-
aINS (R)	-	-	-	**	-	-	-	-	-	***	-	-
SMA	-	-	-	***	-	-	***	-	*	**	*	-

[1]The analysis also included DSMs for “modality”, “cue” and “pattern” (effects not shown).