



38 **Abstract [120 words]**

39

40 Structured sequence processing tasks inform us about statistical learning abilities that are relevant to  
41 many areas of cognition, including language. Despite the ubiquity of these abilities across different  
42 tasks and cognitive domains, recent research in humans has demonstrated that these cognitive  
43 capacities do not represent a single, domain-general system, but are subject to modality- and stimulus-  
44 specific constraints. Sequence processing studies in nonhuman primates have provided initial insights  
45 into the evolution of these abilities. However, few studies have examined similarities and/or  
46 differences in sequence learning across sensory modalities. We review how behavioural and  
47 neuroimaging experiments assess sequence processing abilities across sensory modalities, and how  
48 these tasks could be implemented in nonhuman primates to better understand the evolution of these  
49 cognitive systems.

50

51

52 **Introduction**

53 The ability to recognise and learn predictive dependencies between environmental events is critical to  
54 an animal's survival and is central to a wide range of behaviours. For example, statistical learning—  
55 the development of sensitivity to distributional regularities in an input—appears to be important for  
56 processes as diverse as linguistic processing [1] visual scene analysis [2], motor learning [3] and  
57 many other behaviours that require the prediction of future events [4]. An early suggestion was  
58 therefore that a single cognitive system for extracting statistical regularities might operate over a  
59 number of different domains [5]. In humans, however, direct comparisons across sensory modalities,  
60 or between different types of stimuli, suggest clear modality- and stimulus-specific constraints on how  
61 information is processed [6–8], pointing to differences in the neural systems that underpin these  
62 apparently similar behaviours ([9] and see Fig. 1).

63 Statistical learning experiments, including structured sequence processing tasks and artificial  
64 grammar learning paradigms, can be used to explore the ability to extract order-based regularities  
65 from sequentially-presented stimuli [10,11], (see [12] for a historical review). This approach has  
66 demonstrated that statistical learning abilities likely play a role in language acquisition [1,11] and  
67 syntactic processing [13–15]. Furthermore, comparative experiments have identified similarities in  
68 structured sequence learning across a wide range of nonhuman animals, providing insights into the  
69 types of sequence processing abilities that may have been evolutionary conserved and those which  
70 may have adapted to support language in humans (for reviews, see [16–18]). However, while both  
71 auditory and visual sequence processing have been studied in nonhuman animals, direct comparisons  
72 across modalities are lacking. Such comparisons will be critical in determining how closely the  
73 cognitive systems supporting auditory and visual sequence processing in nonhuman primates  
74 resemble those present in humans.

75 Understanding differences both between the species and across modalities can provide  
76 important insights about potential cognitive specialisations that occurred during more recent human  
77 evolution, and their contributions to the emergence of language. For example, while we might observe  
78 striking similarities in the responses of humans and monkeys using certain stimuli and particular  
79 tasks, it remains possible that very different patterns of learning may be observed across the species  
80 using different stimuli in another modality. Such differences would highlight not only those abilities  
81 that appear to be evolutionarily conserved in nonhuman primates, but might point to behavioural  
82 abilities and the underlying neural substrates which have functionally differentiated in more recent  
83 evolution, and their possible role in language. Identifying such potentially human-unique adaptations  
84 will be critical in understanding how humans diverged from other primates, and how language might  
85 be supported by the human brain [19].

86 In this paper, we summarise how sequence learning has been assessed across sensory  
87 modalities in humans, consider how data from nonhuman animals might be compared in similar ways,

88 and discuss how similarities and differences, across sensory modalities and species, might inform us  
89 about the cognitive and neural systems that support statistical sequence learning.

### 90 **Constraints on sequence processing in humans**

91 A wide range of studies using different stimuli and tasks have shown that humans can extract  
92 statistical regularities from a wide range of sequentially presented auditory or visual stimuli  
93 (summarised in Table 1). These tasks vary in complexity, from learning relatively simple predictive  
94 relationships between adjacent sequence elements, to more nonadjacent or long-distance dependencies  
95 between stimuli, or embedded patterns involving multiple overlapping nonadjacent dependencies (for  
96 reviews see [17,20,21]). However, there is some debate regarding whether statistical learning across  
97 sensory modalities is supported by a single amodal system or by multiple sub-systems that are subject  
98 to stimulus- and modality-specific constraints [9]. While some studies show similar sensitivity to  
99 transitional probabilities between stimuli on matched auditory and visual tasks [22] (see Box 1, Point  
100 1), others report substantial differences. Similarly, although early work identified transfer of learning  
101 from one modality to another [5] (Box 1, Point 2) subsequent studies have suggested that transfer may  
102 be task and structure dependent [23]. In particular, where tasks are based on learning specific  
103 relationships between individual stimuli (e.g. the nonsense word ‘biff’ predicts ‘cav’), transferring the  
104 relationships to a new modality requires learning the mappings between these two stimulus sets, and  
105 therefore is unlikely to occur easily or implicitly. By contrast, more abstract representations or rules  
106 could be more easily transferred between stimuli or modalities as learning is not linked to any specific  
107 stimulus [24], but instead relates to patterns of stimuli (for example element repetitions [23,25]).  
108 Nonetheless in certain tasks information from one modality can influence learning in another (Box 1,  
109 Point 3). For instance, the addition of auditory cues can aid visual sequence learning [26], and  
110 bimodal audio-visual presentation of the same sequence structure results in better performance than  
111 unimodal presentation [27]. However, in humans there is little evidence that individuals’ sequence  
112 learning abilities are correlated across modalities or perceptual domains, further highlighting  
113 stimulus-specific constraints on sequence processing [9,28,29] (Box 1, Point 4). Finally,  
114 neuroimaging work (Box 1, Point 5) can investigate whether the same brain regions are recruited for  
115 sequence learning across modalities. Current evidence paints a complex picture of sequence  
116 processing in the brain (Fig. 1) and is therefore considered in more detail in subsequent sections of  
117 this review. Taken together, this data suggests that there is unlikely to be a unitary sequence  
118 processing mechanism that is tied, for example to general cognitive abilities (for a review see [30]).

### 119 **Sequence learning in primates**

120 In humans, sequence learning is observed reliably across a wide range of tasks and sensory  
121 modalities, albeit with input-related constraints. It is therefore unsurprising that similar learning is  
122 also observed in other species. The study of nonhuman animals, particularly nonhuman primates, has

123 become a valuable way to investigate the evolutionary origins of cognitive and neural systems that  
124 might be related to those that support language in humans [31]. Nonhuman primates have been tested  
125 with a wide variety of different sequence processing tasks [32–37]. Cross-species studies can inform  
126 us about unique adaptations, including specialisations that have been recruited for language in humans  
127 [38], as well as similarities between humans and other primates (see Table 1) [16,22,33,39,40].  
128 Behavioural and neurobiological similarities in sequence learning abilities between humans and other  
129 primate species, suggest that certain sequence processing abilities appear to be evolutionarily  
130 conserved [40–42]. However, there is a lack of evidence about how similarly these systems might  
131 operate across different inputs or sensory modalities, and thus little information as to whether the  
132 variability observed in human sequence learning across different modalities is conserved in nonhuman  
133 animals.

134         In a recent experiment, we directly compared auditory and visual sequence learning in  
135 humans and monkeys [22] (see Box 1, Point 1). This study found similar patterns of responses to a  
136 range of sequences of auditory and visual stimuli, suggesting these processes might be supported by  
137 similar computations [22]. In humans, further insights into the domain-general nature of sequence  
138 processing have been provided by assessing whether learning about one set of stimuli can be  
139 transferred or generalised to novel stimuli or to a different modality (Box 1, Point 2; Table 1).  
140 However, similar experiments have rarely been performed in nonhuman primates. Some studies have  
141 shown that nonhuman primates generalise learning to previously unheard, novel sequences comprised  
142 of familiar stimulus elements [16,32,43], but to date no studies have tested transfer to new stimulus  
143 sets or across modalities. There is some evidence of cross-modal influences, whereby the presentation  
144 of sequences of auditory stimuli might have an impact on visual sequence processing (Box 1, Point 3)  
145 in chimpanzees. In a two-alternative forced-choice experiment, chimpanzees were trained to select  
146 symmetrical rather than asymmetrical sequences of shapes (i.e., XYX vs XYY) [35]. In testing, the  
147 presentation of the visual stimuli was preceded by a previously unheard auditory tone sequence that  
148 was either congruent (symmetrical) or incongruent (asymmetrical) with the visual sequence the  
149 animals were trained to select. The presentation of incongruent auditory stimuli caused an increase in  
150 reaction times, delaying their selection of the appropriate visual sequence [44]. This demonstrates that  
151 properties of the auditory stimuli (i.e., the presence or absence of element repetitions) produced some  
152 interference in visual sequence processing, suggesting at least some cross-modal interactions in great  
153 apes. However, the ability to generalise or transfer statistical regularities has yet to be fully  
154 established in nonhuman primates.

155         In humans there is growing interest in assessing the patterns of individual performance across  
156 sequence learning tasks (Box 1, Point 4; for discussion see [9] and [30]). However, this line of  
157 enquiry has yet to be studied in nonhuman primates. Most primate studies use small sample sizes or  
158 use methods that are hard to replicate in the visual modality [37] - though also see [45]. Although, an

159 opportunity could be provided by recent work in baboons in which voluntary engagement systems  
160 have been shown to produce thousands of trials worth of a data from many animals [46,47].

161 Nonhuman primate research can provide invaluable insights into the evolution and  
162 neurobiology of the systems that support sequence processing. However, in comparative research  
163 there are often unavoidable methodological and cognitive differences between the species which must  
164 be considered [38]. For example, nonhuman primates (and human infants) are often passively exposed  
165 to sequences, while adult humans may be asked to attend to the stimuli, possibly resulting in different  
166 patterns of learning. Similarly, humans can be instructed how to respond, while it is often more  
167 practical to rely on animals' natural orienting responses. Alternatively, animals might be trained using  
168 an operant task for tens of thousands of trials [46,47], making direct comparisons to humans difficult.  
169 There are also unavoidable cognitive differences between humans and other species. Humans may  
170 verbalise or label stimuli, using language to help process stimuli in ways unavailable to nonhuman  
171 primates. They may also try and infer the goal of implicit learning experiments, and respond in the  
172 manner that they think the experimenter desires, which is less likely in nonhuman animals. These  
173 differences must be considered when designing comparative experiments and interpreting their  
174 results, particularly when cross species differences are observed.

175 Nevertheless, the existing behavioural evidence from nonhuman primates indicates that, as in  
176 humans, sequence learning can occur in the auditory and visual modalities, and in primates we  
177 observe similar responses across different types of input [22] as well as some interactions across the  
178 modalities [44]. However, initial human studies also focused on general similarities in statistical  
179 learning. It was only when these capacities were probed in more detail that evidence of modality-  
180 specific constraints on processing emerged. As such, the evidence suggests that humans do not  
181 possess a single, domain-general system that operates identically over all auditory and visual  
182 sequences. Rather the system appears to be more complex and operates under modality and stimulus-  
183 specific constraints. If we are to compare humans and monkeys to draw evolutionary inferences, we  
184 must be careful to compare like to like and not to over-extrapolate from one modality, task, or type of  
185 stimulus to all others. Additional evidence is required to understand if nonhuman primates, like  
186 humans, show sequence learning abilities that vary both qualitatively and quantitatively across  
187 modalities [6], and if these differences were important for the evolution of language.

### 188 **Sequence learning in the brain: across modalities and species**

189 Human neuroimaging experiments using sequence learning and artificial grammar paradigms have  
190 identified a broad network of regions involved in sequence processing (see Fig. 1). Some of these  
191 regions are primarily engaged in only the auditory or visual modality, while other areas are involved  
192 in sequence processing regardless of stimulus modality. In particular, a number of regions such as the  
193 inferior frontal gyrus including the frontal operculum [20] and Broca's territory tend to be engaged by

194 sequence processing tasks in both the auditory [42,48] and visual modality [49,50] (see Fig. 1 and  
195 Table 2). This evidence suggests that overlapping areas are involved in structured sequence learning  
196 across modalities, at least for certain tasks. Importantly, though, some of this overlap might be  
197 attributed to similarities in task demands and response types [20]. For example, comparisons across  
198 tasks that require identification of a violation to the sequence structure (see final column, Table 2)  
199 could reflect similarities in general error detection mechanisms rather than just those which relate to  
200 the extraction of sequence-based regularities.

201         Recently, comparative fMRI experiments using auditory sequence processing tasks in both  
202 humans and macaques [42,43] have demonstrated that sequence violations produced activity in certain  
203 homologous frontal, temporal and parietal regions, particularly inferior frontal regions including the  
204 frontal operculum [43] (see Fig. 1). In this study, activity was also observed in the homologue of  
205 Broca's area in macaques, but not in humans, suggesting potential differentiation of this region [43]  
206 (for a review see [17] and also [42,51]). Visual experiments and direct comparisons across modalities  
207 have yet to be performed using primate neuroimaging, but these will be critical to fully understand the  
208 evolution of the neurobiological systems that support sequence processing (see Fig. 1).

209         While these fMRI studies can provide valuable insights into the neural substrates responsible  
210 for detecting sequence violations, it is also important consider other brain areas within the neural  
211 network involved in sequence processing. Primarily unisensory areas, such as primary auditory and  
212 visual cortex are also likely to play important role in these tasks (Fig. 1 and [2] ) and processing that  
213 occurs within these regions is likely to have implications for operations that occur upstream, in higher  
214 cortical areas (see [9] for a review). In both humans and monkeys, direct recordings of neuronal  
215 responses have highlighted the role of auditory cortex during sequence processing [52]. This study  
216 identified both neurons that showed a preferential response to sequence violations, and others that  
217 responded to sequences that do not contain a violation [52]. These results indicate that even the  
218 earliest cortical regions are sensitive to the order of elements in a sequence (see also [53]). Although  
219 some studies have assessed processing in early visual cortex [2,54], as yet no study has directly  
220 compared how primary auditory and visual cortex respond to identically structured sequences.  
221 Experiments carefully considering the role of sensory cortices and their interactions with other brain  
222 areas including inferior frontal gyrus, either using direct recordings or neuroimaging techniques, are  
223 critical for understanding how different brain regions contribute to the processing of sequence  
224 information, and how this might vary across different stimuli or modalities (Fig. 1).

225

## 226 **Conclusions**

227 Understanding how the brain supports complex cognitive operations, like those involved in sequence  
228 processing, requires rigorous research to differentiate the mechanisms that have been conserved since

229 our last common ancestor with nonhuman primates from those that have diverged. It is initially  
230 tempting to assume that similar patterns in behavioural data point to the presence of a single, domain-  
231 general cognitive or neurobiological system. However, in humans there is little evidence to support  
232 such a conclusion [9]. In primates, there is initial evidence for similar sequence processing abilities,  
233 both between humans and monkeys, and between auditory and visual modalities [22]. However, we  
234 should learn from the human work and not assume that identical processes are at play until we probe  
235 exactly how (and how similarly) auditory and visual sequences are processed, both behaviourally and  
236 in the brain. Another key missing element is the potential role of development in the emergence of  
237 sequence processing skills in nonhuman primates. Our understanding of cross-sensory sequence  
238 processing in nonhuman primates is in its infancy, but by learning from work done in humans, future  
239 research may provide insights that are not possible in humans. These would not only improve our  
240 understanding of how sequence learning abilities evolved, but also the core neuronal computations  
241 and mechanisms which support them.

242

243 **Box 1: Methods of assessing sequence processing across modalities**

244 A number of approaches have been used to assess how sequence processing operates across  
245 different types of stimuli or sensory modalities, to provide insight into the nature of the cognitive  
246 and neural systems involved. These include:

- 247 1. Directly comparing learning of identically structured sequences across different stimuli  
248 or modalities.
- 249 2. Assessing generalisation of learning to new stimuli or transfer to another modality.
- 250 3. Investigating cross-modal influences, such as inhibition or facilitation of the learning of  
251 artificial grammars presented in different modalities.
- 252 4. Exploring correlations in individual performance across statistical learning tasks.
- 253 5. Studying the brain areas and networks engaged in processing sequences presented in  
254 different modalities.

255 Evidence from each of these different approaches can provide important insights into the system(s)  
256 that support sequence learning (see Table 1). However, the data must be carefully considered. For  
257 example, similar patterns of behavioural responses across modalities (e.g., [22]) might be suggestive  
258 of a single, domain-general system. Yet, it is also possible that this result arises from similar  
259 computational principles that are applied in different cognitive or neural systems [9]. Similarly, while  
260 a lack of transfer between modalities suggests some separation in auditory and visual sequence  
261 processing (e.g., [7]), humans may be able to generalise certain stimulus properties (e.g., presence  
262 or absence of repetitions) to novel stimuli, independent of the sequence structure. Evidence of  
263 activation in different brain regions across modalities (e.g., in auditory and visual cortex) can inform  
264 us about the (potentially modality specific) role of initial sensory processing on sequence learning.  
265 However, in cases where both auditory and visual stimuli engage the same brain areas, it is  
266 important to rule out other explanations, such as task-specific effects, before drawing conclusions  
267 about the domain-generalty of the functions of these regions. For example, comparison across tasks  
268 that require identification of a violation to the sequence structure could reflect similarities in general  
269 error detection mechanisms rather than just those which relate to sequence processing. Relatedly,  
270 whether learning and testing occurs in an implicit or explicit paradigm is likely to impact how  
271 different neural systems are engaged [20,55]. Overall, sequence processing is likely supported by  
272 complex cognitive and neurobiological systems (Fig. 1). Understanding the nature of these systems  
273 requires us to carefully consider and interpret the data from several different sources to appreciate  
274 how stimulus- and modality-specific constraints might interact with more domain-general neural  
275 substrates or cognitive computations.

276

277 [Figure 1. attached separately]

278 **Figure 1. Brain areas involved in auditory and visual sequence processing in humans and**  
279 **macaques.** Upper panel (adapted from [9]), shows key brain areas involved in auditory and visual  
280 sequence processing. Brain areas associated with modality-specific auditory and visual processing  
281 are shown in blue and orange circles respectively, and areas involved in domain-general processes in  
282 combined blue and orange circles. These tasks engage a broad network of areas, including areas that  
283 are both primarily unisensory, and those which are involved in both auditory and visual processing. It  
284 may be important to consider the contribution of each of these nodes to fully understand how  
285 sequence processing operates across modalities. This panel illustrates that a broad set of regions are  
286 involved in sequence processing tasks, but that these are not identical across modalities, challenging  
287 the idea of a “domain-general” sequence processing network in the brain. The lower panel shows  
288 the location of anatomical homologues of those regions identified in humans in [9]. Brain areas  
289 involved in auditory [42,43] and visual [56,57] sequence processing tasks in nonhuman primates are  
290 shown in filled blue circles. This highlights that, in the auditory modality similar activation is  
291 observed in humans and monkeys in a number of homologous regions (compare filled and half-filled  
292 blue circles in upper and lower panel), including IFG, STG, IPL and caudate. In monkeys, visual  
293 sequence processing has been measured in inferotemporal cortex using electrophysiological  
294 recordings [56,58] , although other regions are undoubtedly also involved. Therefore, homologues of  
295 the regions seen in visual tasks in humans are denoted by open circles with dashed lines, highlighting  
296 the need for further research into the role of these regions in the visual modality. The depicted  
297 regions are not intended to constitute an exhaustive set of brain regions sub-serving each domain in  
298 either species. Abbreviations: C, cuneus; CA, caudate; FG, fusiform gyrus; H, hippocampus; IFG,  
299 inferior frontal gyrus; IPL, inferior parietal lobule; IT, inferotemporal cortex; STG, superior temporal  
300 gyrus; T, thalamus; A, anterior; P, posterior; D, dorsal; V, ventral; L, left; R, right.

	Experiment	Auditory Stimuli	Visual Stimuli	Artificial Grammar (AG)	Key Results
<b>Humans</b>					
Effects across modalities	<b>Conway and Christiansen, 2005 [6]</b>	Tones	Location	Two Reber-style AGs with probabilistic relationships between adjacent elements	Auditory > visual
	<b>Conway &amp; Christiansen, 2009 [59]</b>	Tones	Textured squares	Reber-style AG with probabilistic relationships between adjacent elements	Fast presentation: Auditory > visual; Slow presentation: Visual > auditory
	<b>Emberson et al., 2011 [8]</b>	Nonsense words	Abstract shapes	Stream segmentation: high probabilities between elements that form 'words' (i.e. triplets of elements), with low probabilities between words	Fast presentation: Auditory > visual; Slow presentation: Visual > auditory
	<b>Walk &amp; Conway, 2016 [29]</b>	Tones/nonsense words	Abstract shapes / colour	Sequences consisting of both auditory and visual stimuli, in which each element could only be followed by one auditory or one visual element	No evidence of cross-modal learning or learning of cross-category dependencies
	<b>Milne et al., 2017 [22]</b>	Sound effects	Abstract shapes	Simplified Reber-style AG with probabilistic relationships between adjacent elements.	Similar patterns of learning across modalities. Visual performance > auditory performance
	<b>Zimmerer et al., 2011 [60]</b>	Syllables	Abstract shapes	A <sup>n</sup> B <sup>n</sup> AG with nonadjacent, embedded relationships between two perceptual classes of stimuli	No significant difference between modalities
Transfer between modalities	<b>Conway and Christiansen, 2006 [7]</b>	Tones	Colours / shapes	Two Reber-style AGs with probabilistic relationships between adjacent elements	Multiple AGs were learned simultaneously if presented in different modalities (no transfer occurred)
	<b>Durrant et al., 2016 [61]</b>	Tones	Location	Deterministic sequences with non-variable relationships between elements	After 24 hours consolidation, deterministic pattern in tones transferred to location of shapes
	<b>Altmann, Dienes &amp; Goode, 1995 [5]</b>	Tones/syllables /nonsense words	Letters/syllables	Two Reber-style AGs with probabilistic relationships between adjacent elements	Transfer from auditory stimuli to visual stimuli, and vice versa.
Cross-modal influences	<b>Mitchel and Weiss, 2011 [62]</b>	Tones	Abstract shapes	Stream segmentation: high probabilities between elements that form 'words' (i.e. triplets of elements), with low probabilities between words	Simultaneous auditory and visual presentation. Learning only occurred in both modalities when statistical boundaries corresponded across modalities
	<b>Mitchel et al., 2014 [63]</b>	Syllables	Abstract shapes	Two Reber-style AGs with probabilistic relationships between adjacent elements	Automatic integration of visual information during auditory statistical learning
	<b>Onnis and Thiessen, 2013 [26]</b>	Italian syllables/tones	Letters	Stream segmentation: high probabilities between elements that form 'words' (i.e. triplets of elements), with low probabilities between words	Visual learning aided by auditory stimuli
	<b>Robinson and Sloutsky, 2007 [64]</b>	Syllables	Shapes and colour	Stream segmentation: high probabilities between elements that form 'words' (i.e. triplets of elements), with low probabilities between words	Statistical information in auditory stream influenced visual learning
	<b>Seitz et al., 2007 [27]</b>	Abstract sounds	Abstract shapes	Stream segmentation: high probabilities between elements that form 'words' (i.e. triplets of elements), with low probabilities between words	Audio-visual sequence learning better than unimodal learning
	<b>van den Bos et al., 2012 [65]</b>	Nonsenses words	Abstract shapes	Probabilistic nonadjacent dependencies	Nonadjacent sequence learning aided by cue from second modality

Correlations across tasks	<b>Siegelman &amp; Frost, 2015 [9]</b>	Syllables/computerised sounds	Abstract shapes	Either deterministic or probabilistic nonadjacent relationships in triplets of elements	No correlations between modalities
<b>Nonhuman primates</b>					
Effects across modalities	<b>Milne et al., 2017 [22]</b>	Sounds effects	Abstract shapes	Simplified Reber-style AG with probabilistic relationships between adjacent elements.	Similar responses across modalities
Transfer between modalities	-	-	-	-	-
Cross-modal influences	<b>Ravignani &amp; Sonnweber, 2017 [44]</b>	Tones	Shapes	Symmetrical vs asymmetrical triplets of elements	Auditory pattern influences visual sequence processing
Correlations across tasks	-	-	-	-	-

**Table 1.** A number of behavioural approaches have been used to assess sequence learning across modalities in humans (top panels) and these are outlined in Box 1 (Points 1 to 4). These include a range of different tasks and the stimuli sequences vary in complexity, assessing the learning of different types of sequencing relationships (for recent reviews, see [17,20]). In humans, these studies provide little evidence for the existence of a single ‘domain general’ sequence processing system, and instead highlight clear stimulus- and modality- specific constraints [9]. Moreover, there does not appear to be a clear link between the types of stimuli or the complexity of the sequences, and cross-modal effects or transfer across modalities. Fewer studies have assessed structured sequence learning across modalities in nonhuman primates (bottom panels). Initial results suggest some similarities across modalities. However, implementing some of the approaches used in human studies in nonhuman primate research will allow us to better understand the constraints under which the sequence processing system(s) operate across modalities and tasks, and how these compare to those observed in humans. This has the potential to provide valuable insights into the evolution of sequence processing abilities, highlighting both those specific abilities and cognitive processes that are evolutionarily conserved, and those which might have diverged and specialised more recently in human evolution.

Modality	Experiment	Stimuli	Artificial grammar	IFG activity	Contrast
<b>Linguistic</b>					
Auditory	<b>Cunillera et al., 2009 [66]</b>	Syllables	Stream segmentation	Left	Sequences/random vs rest
	<b>Goranskaya et al., 2016 [67]</b>	Syllables	A <sup>n</sup> B <sup>n</sup>	None	Learners vs non-learners
	<b>Karuza et al., 2013 [68]</b>	Syllables	Nonadjacent	Left	Forward vs backward order
	<b>Wilson et al., 2015 [43]</b>	Nonsense words	Simplified Reber-style	Bilateral	Violation vs consistent
Visual	<b>Bahlmann et al., 2008 [69]</b>	Syllables	A <sup>n</sup> B <sup>n</sup> vs (AB) <sup>n</sup>	Left	Hierarchical vs adjacent
	<b>Bahlmann et al., 2012 [70]</b>	Syllables	A <sup>n</sup> B <sup>n</sup>	Left	Sequence vs counting
	<b>Folia &amp; Petersson, 2014 [55]</b>	Letters	Reber-style	Bilateral	Violation vs consistent
	<b>Forkstam et al., 2006 [71]</b>	Letters	Reber-style	Left	Classification vs sensorimotor
	<b>Friederici et al., 2006 [49]</b>	Syllables	A <sup>n</sup> B <sup>n</sup> vs (AB) <sup>n</sup>	Left	Violation vs consistent
	<b>Hauser et al. 2012 [72]</b>	Nonsense words	BROCANTO	Right	Consistent vs violation
	<b>Kepinska et al., 2016 [73]</b>	Nonsense words	BROCANTO	Left	Violation vs consistent
	<b>Lieberman et al., 2004 [54]</b>	Letters	Reber-style	Left	Consistent vs violation
<b>Non-Linguistic</b>					
Auditory	<b>Bekinstein et al., 2009 [48]</b>	Tones	Local Global	Bilateral	Global - local violation
	<b>Wang et al., 2015 [42]</b>	Tones	Local Global	Bilateral	Violation vs consistent
Visual	<b>Aizenstein et al., 2004 [74]</b>	Shapes/ colours	Transitional probabilities	Bilateral	Pattern vs no pattern
	<b>Bahlmann et al., 2009 [75]</b>	Abstract shapes	A <sup>n</sup> B <sup>n</sup> vs (AB) <sup>n</sup>	Left	Hierarchical vs adjacent
	<b>Thiel et al., 2003 [76]</b>	Symbols	Bigrams	Bilateral	New vs Old
	<b>van Opstal et al., 2009 [77]</b>	Symbols	Deterministic sequence	Left	Pre-learning vs post-learning

**Table 2.** Summary of fMRI sequence learning studies involving linguistic auditory and visual, and non-linguistic auditory and visual stimuli. Most, but not all, studies showed activity in inferior frontal gyrus (IFG), in Broca's territory and/or the frontal operculum. However, the same artificial grammars are rarely used across modalities, and studies frequently use different contrasts to measure different effects. Furthermore there are relatively few studies that use non-linguistic materials. Direct comparisons using the same artificial grammars across modalities are needed to better understand the neurobiological system that

supports sequence processing. Although a recent meta-analysis highlights the frontal operculum as the region most consistently implicated across artificial grammar learning studies [20].

## **Acknowledgments**

Supported by a Medical Research Council (MRC, U.K.) PhD Studentship to AM; Sir Henry Wellcome Postdoctoral Fellowship to BW (WT110198/Z/15/Z).

## References

1. Saffran JR, Senghas A, Trueswell JC: **The acquisition of language by children.** *Proc Natl Acad Sci U S A* 2001, **98**:12874–5.
2. Turk-Browne NB, Scholl BJ, Chun MM, Johnson MK: **Neural Evidence of Statistical Learning: Efficient Detection of Visual Regularities Without Awareness.** *J Cogn Neurosci* 2009, **21**:1934–1945.
3. Grafton ST, Hazeltine E, Ivry R: **Functional Mapping of Sequence Learning in Normal Humans.** *J Cogn Neurosci* 1995, **7**:497–510.
4. Courville AC, Daw ND, Touretzky DS: **Bayesian theories of conditioning in a changing world.** *Trends Cogn Sci* 2006, **10**:294–300.
5. Altmann G, Dienes Z, Goode A: **Modality independence of implicitly learned grammatical knowledge.** *J Exp Psychol Learn Mem Cogn* 1995, **21**:899.
6. Conway CM, Christiansen MH: **Modality-constrained statistical learning of tactile, visual, and auditory sequences.** *J Exp Psychol Learn Mem Cogn* 2005, **31**:24.
7. Conway CM, Christiansen MH: **Statistical Learning Within and Between Modalities: Pitting Abstract Against Stimulus-Specific Representations.** *Psychol Sci* 2006, **17**:905–912.
8. Emberson L, Conway CM: **Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning.** *Q J* 2011,
9. Frost R, Armstrong BC, Siegelman N, Christiansen MH: **Domain generality versus modality specificity: the paradox of statistical learning.** *Trends Cogn Sci* 2015, **19**:117–125.
10. Reber AS: **Implicit learning of artificial grammars.** *J Verbal Learning Verbal Behav* 1967, **6**:855–863.
11. Saffran JR, Aslin RN, Newport EL: **Statistical Learning by 8-Month-Old Infants.** *Science (80- )* 1996, **274**:1926–1928.
12. Christiansen MH: **Implicit-statistical learning: A tale of two literatures.** *Top Cogn Sci* [date unknown],
13. Conway CM, Pisoni DB: **Neurocognitive basis of implicit learning of sequential structure and its relation to language processing.** *Ann N Y Acad Sci* 2008, **1145**:113–131.
14. Kidd E, Arciuli J: **Individual Differences in Statistical Learning Predict Children’s Comprehension of Syntax.** *Child Dev* 2016, **87**:184–193.
15. Misyak JB, Christiansen MH: **Statistical Learning and Language: An Individual Differences Study.** *Lang Learn* 2012, **62**:302–331.
16. Wilson B, Slater H, Kikuchi Y, Milne AE, Marslen-Wilson WD, Smith K, Petkov CI: **Auditory artificial grammar learning in macaque and marmoset monkeys.** *J Neurosci* 2013, **33**:18825–18835.
17. Wilson B, Marslen-Wilson WD, Petkov CI: **Conserved Sequence Processing in Primate Frontal Cortex.** *Trends Neurosci* 2017, **40**:72–82.
18. Santolin C, Saffran JR: **Constraints on Statistical Learning Across Species.** *Trends Cogn Sci* 2017, doi:10.1016/J.TICS.2017.10.003.
19. Christiansen MH, Chater N: **The language faculty that wasn’t: A usage-based account of natural language recursion.** *Front Psychol* 2015, **6**.
20. Udden, Julia; Männel C: **AGL and its neurobiology.** [date unknown]:In Press.
21. Vries MH De, Christiansen MH, Petersson KM: **Learning Recursion : Multiple Nested and Crossed Dependencies.** *Biolinguistics* 2011,

22. Milne AE, Petkov CI, Wilson B: **Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm.** *Neuroscience* 2017, doi:10.1016/j.neuroscience.2017.06.059.
23. Tunney RJ, Altmann GTM: **The transfer effect in artificial grammar learning: Reappraising the evidence on the transfer of sequential dependencies.** *J Exp Psychol Learn Mem Cogn* 1999, **25**:1322–1333.
24. Daltrozzo J, Conway CM: **Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us?** *Front Hum Neurosci* 2014, **8**:1–22.
25. Gomez RL, Gerken L, Schvaneveldt RW: **The basis of transfer in artificial grammar learning.** *Mem Cognit* 2000, **28**:253–263.
26. Onnis L, Thiessen E: **Language experience changes subsequent learning.** *Cognition* 2013, **126**:268–284.
27. Seitz AR, Kim R, van Wassenhove V, Shams L: **Simultaneous and independent acquisition of multisensory and unisensory associations.** *Perception-London* 2007, **36**:1445–1454.
28. Cope TE, Wilson B, Robson H, Drinkall R, Dean L, Grube M, Jones PS, Patterson K, Griffiths TD, Rowe JB, et al.: **Artificial grammar learning in vascular and progressive non-fluent aphasias.** *Neuropsychologia* 2017, doi:10.1016/j.neuropsychologia.2017.08.022.
29. Walk AM, Conway CM: **Cross-Domain Statistical–Sequential Dependencies Are Difficult to Learn.** *Front Psychol* 2016, **7**:1–9.
30. Siegelman N, Bogaerts L, Christiansen MH, Frost R: **Towards a theory of individual differences in statistical learning.** *Philos Trans R Soc B Biol Sci* 2017, **372**:20160059.
31. Fitch WT: *The evolution of language.* Cambridge University Press; 2010.
32. Hauser MD, Glynn D: **Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations?** *J Comp Psychol* 2009, **123**:161.
33. Newport EL, Hauser MD, Spaepen G, Aslin RN: **Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate.** *Cogn Psychol* 2004, **49**:85–117.
34. Ravignani A, Sonnweber R-S, Stobbe N, Fitch WT: **Action at a distance: dependency sensitivity in a New World primate.** *Biol Lett* 2013, **9**:20130852.
35. Sonnweber R, Ravignani A, Fitch WT: **Non-adjacent visual dependency learning in chimpanzees.** *Anim Cogn* 2015, **18**:733–745.
36. Heimbauer LA, Conway CM, Christiansen MH, Beran MJ, Owren MJ: **A Serial Reaction Time (SRT) task with symmetrical joystick responding for nonhuman primates.** *Behav Res Methods* 2012, **44**:733–741.
37. Saffran JR, Hauser MD, Seibel RL, Kapfhamer J, Tsao F, Cushman F: **Grammatical pattern learning by human infants and cotton-top tamarin monkeys.** *Cognition* 2008, **107**:479–500.
38. Conway CM, Christiansen MH: **Review: Sequential learning in non-human primates.** *Trends Cogn Sci* 2001, **5**:539–546.
39. Hauser MD, Newport EL, Aslin RN: **Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins.** *Cognition* 2001, **78**:B53–B64.
40. Attaheri A, Kikuchi Y, Milne AE, Wilson B, Alter K, Petkov CI: **EEG potentials associated with artificial grammar learning in the primate brain.** *Brain Lang* 2015, **148**:74–80.
41. Milne AE, Mueller JL, Männel C, Attaheri A, Friederici AD, Petkov CI: **Evolutionary origins of non-adjacent sequence processing in primate brain potentials.** *Sci Rep* 2016, **6**:36259.
42. Wang L, Uhrig L, Jarraya B, Dehaene S: **Representation of numerical and sequential**

- patterns in macaque and human brains.** *Curr Biol* 2015, **25**:1966–1974.
43. Wilson B, Kikuchi Y, Sun L, Hunter D, Dick F, Smith K, Thiele A, Griffiths TD, Marslen-Wilson WD, Petkov CI: **Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans.** *Nat Commun* 2015, **6**:1–19.
  44. Ravnani A, Sonnweber R: **Chimpanzees process structural isomorphisms across sensory modalities.** *Cognition* 2017, **161**:74–79.
  45. Siegelman, N., Bogaerts, L., Kronenfeld, O. & Frost R: ). **Re-defining “learning” in statistical learning: what does an online measure reveal about the assimilation of visual regularities?** *Cogn Sci* [date unknown],
  46. Fagot J, Gullstrand J, Kemp C, Defilles C, Mekaouche M: **Effects of freely accessible computerized test systems on the spontaneous behaviors and stress level of Guinea baboons (*Papio papio*).** *Am J Primatol* 2014, **76**:56–64.
  47. Grainger J, Dufau S, Montant M, Ziegler JC, Fagot J: **Orthographic Processing in Baboons (*Papio papio*).** *Science (80- )* 2012, **336**:245–248.
  48. Bekinschtein TA, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L: **Neural signature of the conscious processing of auditory regularities.** *Proc Natl Acad Sci* 2009, **106**:1672–1677.
  49. Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A: **The brain differentiates human and non-human grammars: Functional localization and structural connectivity.** *Proc Natl Acad Sci U S A* 2006, **103**:2458–2463.
  50. Petersson KM, Forkstam C, Ingvar M: **Artificial syntactic violations activate Broca’s region.** *Cogn Sci* 2004, **28**:383–407.
  51. Christiansen, Morten H, Mueller RA: **Cultural recycling of neural substrates during language evolution and development.** In *The Cognitive Neurosciences V*. Edited by Gazzaniga MS, Mangun GR. Cambridge, MA: MIT Press; 2014:675–682.
  52. Kikuchi Y, Attaheri A, Wilson B, Rhone AE, Nourski K V, Gander PE, Kovach CK, Kawasaki H, Griffiths TD, Howard MA, et al.: **Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex.** *PLOS Biol* 2017, **15**:e2000219.
  53. Hasson U: **The neurobiology of uncertainty: implications for statistical learning.** *Philos Trans R Soc B Biol Sci* 2017, **372**:20160048.
  54. Lieberman MD, Chang GY, Chiao J, Bookheimer SY, Knowlton BJ: **An Event-related fMRI Study of Artificial Grammar Learning in a Balanced Chunk Strength Design.** *J Cogn Neurosci* 2004, **16**:427–438.
  55. Folia V, Petersson KM: **Implicit structured sequence learning: An fMRI study of the structural mere-exposure effect.** *Front Psychol* 2014, **5**:1–13.
  56. Meyer T, Olson CR: **Statistical learning of visual transitions in monkey inferotemporal cortex.** *Proc Natl Acad Sci* 2011, **108**:19401–19406.
  57. Meyer T, Olson CR: **Statistical learning of visual transitions in monkey inferotemporal cortex.** *Proc Natl Acad Sci* 2011, **108**:19401–19406.
  58. Meyer T, Ramachandran S, Olson CR: **Statistical Learning of Serial Visual Transitions by Neurons in Monkey Inferotemporal Cortex.** *J Neurosci* 2014, **34**:9332–9337.
  59. Conway CM, Christiansen MH: **Seeing and hearing in space and time: Effects of modality and presentation rate on implicit statistical learning.** *Eur J Cogn Psychol* 2009, **21**:561–580.
  60. Zimmerer VC, Cowell PE, Varley RA: **Individual behavior in learning of an artificial grammar.** *Mem Cognit* 2011, **39**:491–501.
  61. Durrant SJ, Cairney SA, Lewis PA: **Cross-modal transfer of statistical information benefits**

- from sleep. *Cortex* 2016, **78**:85–99.
62. Mitchel AD, Weiss DJ: **Learning across senses: Cross-modal effects in multisensory statistical learning.** *J Exp Psychol Learn Mem Cogn* 2011, **37**:1081–1091.
63. Mitchel AD, Christiansen MH, Weiss DJ: **Multimodal integration in statistical learning: Evidence from the McGurk illusion.** *Front Psychol* 2014, **5**.
64. Robinson CW, Sloutsky VM: **Visual processing speed: effects of auditory input on visual processing.** *Dev Sci* 2007, **10**:734–740.
65. van den Bos E, Christiansen MH, Misyak JB: **Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration.** *J Mem Lang* 2012, doi:10.1016/j.jml.2012.07.008.
66. Cunillera T, Càmara E, Toro JM, Marco-Pallares J, Sebastián-Galles N, Ortiz H, Pujol J, Rodríguez-Fornells A: **Time course and functional neuroanatomy of speech segmentation in adults.** *Neuroimage* 2009, **48**:541–553.
67. Goranskaya D, Kreitewolf J, Mueller JL, Friederici AD, Hartwigsen G: **Fronto-Parietal Contributions to Phonological Processes in Successful Artificial Grammar Learning.** *Front Hum Neurosci* 2016, **10**.
68. Karuza EA, Newport EL, Aslin RN, Starling SJ, Tivarus ME, Bavelier D: **The neural correlates of statistical learning in a word segmentation task: An fMRI study.** *Brain Lang* 2013, **127**:46–54.
69. Bahlmann J, Schubotz RI, Friederici AD: **Hierarchical artificial grammar processing engages Broca's area.** *Neuroimage* 2008, **42**:525–534.
70. Bahlmann J, Korb FM, Gratton C, Friederici AD: **Levels of Integration in Cognitive Control and Sequence Processing in the Prefrontal Cortex.** *PLoS One* 2012, **7**:e43774.
71. Forkstam C, Hagoort P, Fernández G, Ingvar M, Petersson KM: **Neural correlates of artificial syntactic structure classification.** *Neuroimage* 2006, **32**:956–967.
72. Hauser MFA, Hofmann J, Opitz B: **Rule and similarity in grammar: Their interplay and individual differences in the brain.** *Neuroimage* 2012, **60**:2019–2026.
73. Kepinska O, Rover M de, Caspers J, Schiller N: **On neural correlates of individual differences in novel grammar learning: an fMRI study.** *Neuropsychologia* 2017,
74. Aizenstein HJ, Stenger VA, Cochran J, Clark K, Johnson M, Nebes RD, Carter CS: **Regional Brain Activation during Concurrent Implicit and Explicit Sequence Learning.** *Cereb Cortex* 2004, **14**:199–208.
75. Bahlmann J, Schubotz RI, Mueller JL, Koester D, Friederici AD: **Neural circuits of hierarchical visuo-spatial sequence processing.** *Brain Res* 2009, **1298**:161–170.
76. Thiel CM, Shanks D, Henson R, Dolan R: **Neuronal correlates of familiarity-driven decisions in artificial grammar learning.** *Neuroreport* 2003, **14**:131–136.
77. Van Opstal F, Fias W, Peigneux P, Verguts T: **The neural representation of extensively trained ordered sequences.** *Neuroimage* 2009, **47**:367–375.

## Annotations:

**\*Christiansen and Chater, 2015.** Suggest that the ability to process recursive structures in language derives from complex sequence learning skills evolved in the human lineage. Constraints on sequence learning is argues to have played an important role in the cultural evolution of linguistic structure, including the limited ability to process recursive constructions.

**\*Durrant et al., 2016.** Using a simple statistical learning task with deterministic sequences, transfer was observed from the auditory to the visual modality but only after a 24-hour consolidation period. Initial evidence is provided showing the relevance of consolidation for cross-modal transfer that requires further investigation using probabilistically structured sequences.

**\*Milne et al., 2017.** In the first study to directly test structure sequence learning abilities across species (human vs. macaque) and modalities (auditory vs. visual), the same artificial grammar was used to generate sequences of computer-generated sound effects or abstract shapes. Both species were sensitive to violations of the artificial grammar and showed patterns of responses were highly consistent across the two modalities. These data suggest that similar computations are likely to occur across modalities in the both human and nonhuman primates.

**\*Siegelman & Frost, 2015.** Human participants were tested on a range of statistical learning tasks using auditory and visual, verbal and non-verbal stimuli. The results found that performance was not correlated across the tasks showing that at an individual level statistical sequence learning abilities do not reflect a unified capacity.

**\*Walk and Conway, 2016.** In a multimodal sequence learning experiment subjects could not learn relationships between items of different perceptual categories or perceptual modalities. This study demonstrates that statistical learning can operate within but not across domains.

**\*Wilson et al., 2015.** Comparative fMRI was used to identify key brain areas in ventral frontal cortex which are similarly involved in auditory sequence processing in both macaque monkeys and human participants. In humans, this region is plays a role in syntactic processing. These results identify evolutionarily conserved neural substrates that are involved in sequence processing.