

# Hierarchical processing in music, language, and action: Lashley revisited

W. Tecumseh Fitch<sup>1</sup> and Mauricio D. Martins<sup>1,2</sup>

<sup>1</sup>Department of Cognitive Biology, University of Vienna, Vienna, Austria <sup>2</sup>Language Research Laboratory, Lisbon Faculty of Medicine, Lisbon, Portugal

Address for correspondence: W. Tecumseh Fitch, Department of Cognitive Biology, University of Vienna, 14 Althanstrasse A-1090, Vienna, Austria. tecumseh.fitch@univie.ac.at

Sixty years ago, Karl Lashley suggested that complex action sequences, from simple motor acts to language and music, are a fundamental but neglected aspect of neural function. Lashley demonstrated the inadequacy of then-standard models of associative chaining, positing a more flexible and generalized “syntax of action” necessary to encompass key aspects of language and music. He suggested that hierarchy in language and music builds upon a more basic sequential action system, and provided several concrete hypotheses about the nature of this system. Here, we review a diverse set of modern data concerning musical, linguistic, and other action processing, finding them largely consistent with an updated neuroanatomical version of Lashley’s hypotheses. In particular, the lateral premotor cortex, including Broca’s area, plays important roles in hierarchical processing in language, music, and at least some action sequences. Although the precise computational function of the lateral prefrontal regions in action syntax remains debated, Lashley’s notion—that this cortical region implements a working-memory buffer or stack scannable by posterior and subcortical brain regions—is consistent with considerable experimental data.

**Keywords:** hierarchy; language; music; computation; syntax

## Introduction

In 1951, Karl Lashley<sup>1</sup> suggested that the human capacity for serial ordering of action, whether of language or music or simply making coffee, presents central problems for our understanding of brain function. Lashley observed that then-prevailing models of action based on simple stimulus–response chaining had a fundamental flaw: although such serial models can capture the “one thing after another” characteristic of routine action, they fail to recognize the central importance of sustained goals and subgoals in more complex action. In complex actions, overall goals (e.g., making coffee) must persist over time while subgoals are initiated and completed (e.g., grinding beans, heating water, and adding cream) or interruptions occur (e.g., a phone call). Several aspects of such action sequences are inconsistent with a chaining model. Most obvious are errors of omission, where some step is accidentally left out (e.g., starting the coffee machine without first adding coffee). In chaining models, each

completed action serves as the stimulus for the next one, and omission of any step would cause the sequence to grind to a halt. This and other phenomena, argued Lashley, demonstrate the need for hierarchical models of action planning—a requirement widely acknowledged today.<sup>2</sup>

Lashley further observed that core aspects of spoken language and music share this need for hierarchical structure, at multiple levels. For example, when pronouncing the words *tire* and *right*, the same basic vocal gestures appear but in different orders. How could a chaining system cope with such flexible rearrangement of basic actions into multiple larger composites? Phrasal syntax makes the point even more clearly. When reading the sentence, “The boy who patted the dog chased the girl,” an English speaker knows that the boy, and not the dog, chased the girl, despite the fact that this sentence contains the sequence “the dog chased the girl.” Linguistic syntax illustrates that serial order alone provides an inadequate basis for language processing, and that processes involving longer temporal

scales must persist during the processing or production of shorter actions or simple chains. Such observations were foundational in early generative linguistics, which recognized the central role of phrase structure, rather than word order, in syntax.<sup>3,4</sup>

Lashley also suggested that these similarities in language, music, and other complex actions are not coincidental, instead hypothesizing that the hierarchical nature of music and language was inherited from more basic features of motor planning and action.

In this review, we reassess Lashley's observations and hypotheses from a modern perspective, concluding that his ideas were remarkably insightful and consistent with considerable new behavioral and neuroscientific data. We explore Lashley's idea that hierarchical structuring of temporal sequences is a key capability underlying human music and language, but one that has deeper phylogenetic roots in the domain of action. Unlike Lashley, who eschewed localizationist arguments, we make full use of brain imaging and lesion data as supporting evidence, with a particular focus on the role of Broca's area in processing hierarchical structures in the time domain.

Specifically, we will consider the following issues: (1) whether processing of sequential hierarchy is a well-defined specific capability, independent of other hierarchical processing (e.g., visuospatial) but applicable to both music and language; (2) to what extent the mechanisms subserving linguistic and musical syntax inherit properties from those typifying a syntax of action more generally; and (3) whether humans are unusual in the degree to which hierarchical abilities are developed, and if so, which neural mechanisms support this proclivity.

We evaluate these questions using behavioral, neuroanatomical, and brain-imaging data that go well beyond the data available in Lashley's day. We conclude that these data converge upon the notion of a distinct neural substrate for hierarchical sequence processing, unusually well developed in humans relative to other primates, and reflected in specific expansions of prefrontal brain regions including Broca's area, as well as a considerable remodeling of connections between prefrontal and posterior regions. Although the degree to which music and language inherit key properties from action syntax remains controversial, available data are consistent with overlap in these systems from both

neural and computational viewpoints. More precise computational models and experimental designs will be required to further evaluate this hypothesis, which is quite different from currently popular hypotheses based upon mirror neurons and embodiment. Thus, 60 years on, Lashley's insights remain valid and intriguing and can inform current neuroscientific debates, and thus deserve renewed attention.

### Defining temporal hierarchy: hierarchical sequences and hierarchical sets

Terms like *syntax* and *hierarchy* are notoriously ambiguous, so our first task in evaluating Lashley's hypotheses is to distinguish possible interpretations of the term hierarchy, which has many meanings in neuroscience. Especially in older literature, hierarchy connotes little more than the fact that cortical regions are above or below one another in processing terms, such that bottom-up processing proceeds from regions lower to regions higher in the visual hierarchy. A central characteristic of the visual system is that neurons in cortical regions closer to retinal input, such as V1, have smaller receptive fields than those in higher regions like V2. Our focus goes beyond this basic and widespread form of hierarchical layout of cortical regions themselves.

Here, we adopt a more specific and precise notion of hierarchy to denote a tree-like organization, where higher levels incorporate multiple lower levels in structural representations and/or processing. This notion is wholly independent of how the cortical regions that process such stimuli are arranged, entailing no specific commitments about neuroanatomical localization.

To be more exact we utilize the following mathematically-based definitions<sup>5,6</sup> throughout this paper: a *set* is an unordered collection of distinct, unique objects; while a *sequence* is a collection of objects, perhaps including duplicates, ordered by some rule.

Although the set {a, b, c} is identical to the set {b, c, a}, the sequences [a b c] and [b c a] are distinct and different. Furthermore, because sequences but not sets can contain duplicate items, sequences are not, strictly speaking, a type of set. These distinctions are crucial in distinguishing Lashley's temporally ordered sequences, where order matters, from static hierarchy, in which order is typically irrelevant.

*Hierarchy* denotes a set or sequence of elements connected in the form of a rooted tree (a connected acyclic graph, in which one element is singled out as the root element). Hierarchies thus possess the following key properties: (1) all elements are combined into one structure (connectedness); (2) one element (the root) is superior to all others; and (3) no element is superior to itself (that is, there are no cycles, direct or indirect).

These definitions allow us to clarify a core distinction between temporally ordered and unordered hierarchies. Examples of hierarchical sets include visuospatial part-whole hierarchies, such as the schema of a face, in which certain components contain others without any necessary sequential ordering of components at each level. A face incorporates the set {eyes, nose, mouth}, but there is no need to consider one component to come first. Such hierarchies involve the minimal structure required to consider a representation hierarchical, namely that the superior/inferior relation between levels is specified. But at any given level, the elements form an unordered set.

In hierarchical sequences, in contrast, sequential order matters. The simple two-syllable utterances “callow” and “low-cal” contain the same set of syllables, but arranged in different orders, and they represent different concepts. Similarly, in music each arrangement of a set of notes constitutes a different melody. Finally, in action sequences, order is often important: if we grind the coffee beans at the end of the action sequence, after brewing is completed, poor coffee results. Thus temporal hierarchies incorporate an additional ordering component, where at least some elements at any given level represent a sequence rather than a set.

Lashley considered hierarchical structure to be an obvious component of language, music, and action. In each domain, some superordinate structure (his “determining tendency”) is required to lend coherence to subordinate components (speech sounds, notes, or simple actions). Although these higher level elements may differ considerably between domains (e.g., meaning for language versus an overall goal for action), Lashley’s focus was what all three domains have in common: the need to impose the correct temporal ordering on the subelements. He cites, as a key example, language, where the elements of meaning are often cotemporal sets, but different languages have different ordering rules by which

the same meaning must be expressed in a sequence: different “schemas of order.” Typing or speech errors (e.g., spoonerisms) or children’s games like pig Latin provide evidence that the temporal ordering component is both independent of semantics, and a general characteristic of many “generalized schemata of action” that apply across the three domains. Thus, it is not the general problem of hierarchy but the specific issue of hierarchical sequencing that Lashley considered the essential problem of serial order, and that will be our focus in this review.

Lashley observed that we can often convert between spatial and temporal hierarchies. For instance, to reverse a piano melody, he stated, “I can only do it by visualizing the music spatially and then reading it backward” (p. 129). Similarly, when translating between languages with different word orders, such as English and German, a fluent bilingual can extract the meaning from one language, and convert it to a properly sequenced sentence in the other, freely reordering the individual words as needed. It is therefore possible that the distinction between sequential and nonordered hierarchy is of little importance at a fundamental neural or computational level. This is the first issue that we address using modern imaging data.

### **Sequential hierarchical processing is not colocalized with visual hierarchy systems**

As noted above, there is often an increase in spatial receptive field size in cortical regions more distant from sensory input; this involves a widespread neural form of spatial hierarchy, where higher regions integrate information from multiple lower regions.<sup>7</sup> In this functional sense, hierarchical processing is a widespread property of brain systems.<sup>8–16</sup>

In some cases, such patterns of neuroanatomical layout may reflect a functional form of cognitive hierarchy. For instance, in the motor domain, posterior frontal lobe regions encode basic movements (extending the arm), whereas anterior regions fire more selectively to particular contexts in which these movements are performed (e.g., making coffee and shaking hands).<sup>10</sup> In the auditory cortex, core areas respond to simple acoustic characteristics, whereas anterior regions along the supratemporal plane display selective firing to more complex sound categories.<sup>12</sup> In the visual ventral stream, anterior regions in the temporal cortex support

categorical decisions of greater abstraction (e.g., animal versus nonanimal) than posterior regions (e.g., blue versus green).<sup>14</sup> Given these pervasive patterns of brain organization, is there evidence for more abstract and modality independent hierarchical structuring in the brain?

Starting with hierarchical sets, many recent studies explore the brain regions involved in representing and processing hierarchical structures in both spatial and social domains.<sup>17,18</sup> Representation of hierarchical structures in the social domain, which allows the evaluation of dominance relationships, appear to be encoded in the hippocampus.<sup>18</sup> However, this structure is also active in encoding hierarchical ranks in nonsocial domains.<sup>18</sup> In the visuospatial domain, for example, correct integration of landmarks within context frames recruits the parahippocampus,<sup>8</sup> and also requires the integrity of the medial temporal lobe (MTL) and the retrosplenial cortex.<sup>17</sup>

Interestingly, this same MTL system may also encode high order hierarchical associations in the motor and linguistic domains.<sup>19–22</sup> It has been proposed that these medial structures are important for binding items within contexts,<sup>23,24</sup> even for offline imagined scenes,<sup>25</sup> as well as for episodic memory, which involves binding different perceptual and memory features into unified representations.<sup>26,27</sup>

In addition to these brain regions, a quite different set of activations is typically observed in temporal hierarchical processing, regardless of whether these involve musical, linguistic, or other tasks. A consistent finding within the neurolinguistic literature is that sequentially structured hierarchical processing activates the posterior prefrontal cortex, centering on Broca's area (a term we will use hereafter to denote both Brodmann's areas [BA] 44 and 45, typically in the left hemisphere, but not neighboring areas<sup>28</sup>). Increased Broca's activation typically correlates with increased demands on working memory, for example, when processing long-distance dependencies, where early information must remain active for proper interpretation of later elements. Ongoing controversy concerns the degree to which this active maintenance of past information is specific to hierarchical sequences per se<sup>22,29,30</sup> or is found for nonhierarchical sequences as well,<sup>31,32</sup> but it is clear that increasingly complex hierarchical structures lead to greater Broca's activation across multiple languages and laboratories.

Furthermore, in musical syntax, hierarchically dependent processing can stretch over long time spans (long-distance dependency).<sup>33</sup> Neuroimaging studies on musical syntax (as probed with sequential harmonic or melodic structures) also consistently reveal activation in BA 44/45, typically in both the right and left hemispheres, but sometimes biased to the right side.<sup>34–39</sup> Evidence that these prefrontal activations are not restricted to meaningful language or music comes from artificial grammar research using nonsense syllables (spoken) or words (written) arranged according to particular rules. Again, irrespective of the input domain, both auditory and visual patterns, and even nonlinguistic visual symbols,<sup>40</sup> elicit Broca's activation when they are sequentially presented. On the basis of these data, further explored below, we conclude that, despite MTL overlaps, the neural mechanisms for processing hierarchical sequences are not fully colocalized with those that process hierarchical sets. Despite formal similarities in these abstract hierarchical structures, these data argue against their identity. We now explore the functions of the prefrontal regions preferentially involved in hierarchical sequence processing.

### **Prefrontal cortex and Broca's area play central roles in processing hierarchical syntax**

We first attempt to characterize the capacity for processing sequential hierarchies more precisely, to better understand its neural basis. Although Lashley was a confirmed agnostic regarding neural localization, there is a long tradition of associating the frontal lobes, specifically the prefrontal cortex (the portion of the frontal lobe anterior to the primary motor strip and the premotor cortex), with the organization of complex plans in general, and hierarchical sequences in particular, in multiple cognitive domains. The notion that the prefrontal cortex is specifically involved in action planning and executive control has been advanced notably by Luria,<sup>41</sup> Shallice,<sup>42</sup> and Baddeley,<sup>43</sup> among many others. Clinical data leave little doubt that pathology in the frontal lobe can lead to disorganized action.<sup>44</sup> We do not review this extensive literature here, but simply take it for granted that the frontal lobes play important roles in planning, working memory, and executive control. Our goal will be to specify this contribution more precisely, paying particular

attention to the role of the inferior frontal gyrus (IFG) and Broca's area.

We note first that virtually any motor behavior can be conceptualized as hierarchical. A monkey eating peanuts involves a set of actions (grasp, ingest, chew) each of which can be further broken down (grasp: extend arm, open fingers, close fingers, retract arm). Thus, any action sequence that involves repeated subparts displays a simple form of hierarchy, pervasive to all motor action in all species.<sup>45</sup> Such hierarchies have a fixed and limited depth, and were not Lashley's focus (because they can be accounted for by chaining). To clarify what he was getting at, we start with the clear case of language, and use this to guide our further inquiry through music, back to action.

Broca's area (BA 44 and 45 or, roughly, the pars opercularis and the pars triangularis of the IFG) is one of the most intensively studied brain regions, and its importance for speech production requires little emphasis.<sup>46–48</sup> Since the 1970s, this region has also been recognized in clinical studies to play a special role in syntax, both production and perception.<sup>49</sup> Thus, a broad consensus exists that this region of the IFG plays an important role in language processing.<sup>30,50–52</sup>

An elegant study by Pallier *et al.*<sup>53</sup> explored the specificity of IFG involvement in hierarchical syntax by crossing two manipulations. In the first, they parametrically varied the chunk size of 12-word sentences, spanning from scrambled words (where there are no chunks, giving a constituent size of one) to full sentences (constituent size = 12). Intermediate strings had constituent sizes of two, three, four, and six words (an example with four-word chunks would be “mayor of the city/he hates this color/they read their names”). The prediction was that activation in a region specifically encoding constituent structure would increase with greater constituent size. Such a correlation was indeed found, including both IFG structures and temporal lobe areas, especially along the superior temporal sulcus.

To get at syntax specificity, a second manipulation involved substituting all content words with nonsense words to create “jabberwocky” constituents of varying sizes.<sup>53</sup> For example, the four-word jabberwocky could read “tuyor of the roty/he futes this dator/they gead their wames.” This parametric manipulation removes the effect of semantic integration, but again elicited specific correlated activation in the

IFG (centered on BA 45). The clever design controls for both semantics and word transition probability (because the nonsense words, by definition, have zero probability), and confirmed a specific role for Broca's area in processing hierarchical constituent structure in language.

### Broca's region is involved in processing musical harmonic syntax

Regarding specificity of function, however, there has been considerable debate about the degree to which the function of Broca's area is specific to syntax, or whether it performs one or more computational functions that are shared by multiple aspects of language, including phonology or semantics,<sup>50</sup> or even more broad cognition including cognitive control, action planning, and music.<sup>35,54–56</sup>

The best characterized, and to us most convincing, nonlinguistic function of the IFG is in music processing and memory.<sup>57–60</sup> There is a long tradition, based on dissociations in lesion patients, of considering music and language to be instantiated in distinct brain regions.<sup>61</sup> Initial evidence for a shared role for the IFG in musical processing, particularly processing harmonic syntax, came from event-related potential (ERP) and magnetoencephalography (MEG) studies.<sup>36,37,62</sup> Furthermore, there are multiple studies indicating overlap of both cognitive and neural resources used in music-syntactic and language-syntactic processing based on interactions between the two domains.<sup>63–65</sup> These studies showed consistent similarities in early, anteriorly located reactions to anomalous items in linguistic and musical stimuli, and the MEG study by Maess *et al.*<sup>37</sup> concluded simply that “musical syntax is processed in Broca's area.”

These results, from normal subjects, stand in contrast to the long clinical tradition identifying patients with damage to music and not language, or vice versa.<sup>61</sup> Recent commentators<sup>57,58</sup> have reconciled this apparent discrepancy by noting that although certain abstract operations may be shared by music and language, the units over which these functions operate are clearly quite different (phonemes and words for language, notes and chords for music). Furthermore, there are fundamental differences among different components of music. For example, a key component of musical rhythm is *isochronicity*—the tendency of key elements to be equally spaced in time.<sup>66</sup> This is not an aspect of

rhythm shared by normal speech or action, although other aspects like metrical structure may represent hierarchical sequencing.<sup>67</sup> Given these basic differences, we do not expect complete overlap of musical and linguistic processing, either neurally or behaviorally. Rather, specific parallels in the manipulations applied to musical and linguistic stimuli (e.g., violations of hierarchical structural expectations<sup>33</sup>) are needed to elicit parallel neural effects.

Despite the broad consensus concerning a role for the IFG in these aspects of music processing, both ERP and MEG studies have difficulties with localization, and these early findings have not gone unchallenged by more recent functional magnetic resonance imaging (fMRI) studies.<sup>68,69</sup> A particular issue concerns group analysis of different subjects. It is possible that each individual in a study could activate distinct but neighboring regions for linguistic versus musical syntax. Because of individual variability in precise localization, averaging all these individually distinct activations would give a false impression of thorough overlap in the two domains.

A recent study probed this issue by first identifying language regions in individual subjects, and then examining activations elicited in those regions by various other tasks involving working memory, cognitive control, and music.<sup>70</sup> This study used a sentence-reading task, contrasted with reading nonsense words, to delineate the language regions of interest (ROIs). Although the study found considerable specificity to language in IFG activations, a musical task also preferentially activated their language-based ROIs, in particular overlapping with activations found in the left posterior IFG for both the sentence-reading task and a verbal working-memory task. Although the authors concluded from their results that specific language areas exist, not activated by most other tasks, this is an overly general conclusion applicable only to their specific musical and linguistic tasks. It is also important to recognize that the Patel or Koelsch shared-resource models<sup>57,58</sup> do not predict exact overlap for these two different tasks (reading visually presented sentences versus listening to melodies) given the different input modalities. Thus, the significant language-area activations that were seen for the musical task support Lashley's contention that these two domains share important processing subsystems.

A second source of apparently contradictory evidence involves rhythm perception. A series of re-

cent imaging studies concerning rhythm perception paint a consistent picture:<sup>71–73</sup> rhythmic stimuli tend to activate motor areas (such as the supplementary motor area, other dorsal prefrontal regions, or even the basal ganglia) involved in sequencing,<sup>74,75</sup> and do not activate the IFG. But this work focuses on isochronous rhythms, and likely reflects the tendency of such rhythms to induce a desire to move to the regular beat—the so-called “groove” of the music.<sup>76,77</sup> Precisely because this aspect of music is not shared with language, we do not consider such data to contradict Lashley's hypothesis or the well-established notion that the IFG plays a role in processing other components of musical syntax.

In summary, a considerable body of neuroimaging data, from many laboratories, is consistent with the hypothesis that the prefrontal cortex, and the IFG in particular, play important parallel roles in processing specific shared aspects of language and music. These data are consistent with both Lashley's early ideas and Patel and Koelsch's more specific hypotheses about the nature of the shared processing resources.

### **The role of frontocortical mirror neurons in understanding action**

The third component of Lashley's triad is goal-directed action. There has been an explosion of interest recently in the role of the frontal lobes, and the IFG, in action understanding. This work typically makes no reference to Lashley's ideas, but was instead spurred by the discovery of *mirror neurons* in the macaque. Mirror neurons are neurons in motor and premotor cortex that fire, not only when a monkey executes an action itself, but also when it sees another agent perform the same action. The dual nature of the metaphoric mirror—in both production and perception—has led to revived interest in the mostly discredited motor theory of speech perception, which posits that speech perception is mediated by virtue of a resonance with the motor programs that generate the sounds (vocal gestures).<sup>78</sup> By this hypothesis, the objects of speech perception are articulatory motor events, rather than acoustic or auditory events (for a critique, see Ref. 79).

Recent speculation concerning mirror neurons has extended the reach of the motor theory beyond its original domain of speech perception to language more generally. The region of macaque cortex where mirror neurons are found—F5—neighbors or

overlaps the regions constituting the human Broca's area (BA 44/45). Many studies show activation of motor and premotor regions during language production and comprehension tasks,<sup>80,81</sup> thought to support the notion that mirror neurons in this region play a central role in the parity of production and perception that typifies all aspects of language.<sup>82–84</sup> This idea has been used to support gestural hypotheses for language evolution, whereby initial stages of linguistic syntax and semantics evolved in the context of a gestural protolanguage, rather than a vocal system.<sup>83,85–87</sup>

A detailed critique of mirror neurons or their possible importance in language evolution is not our purpose here (but see Refs. 88–92). However, it is crucial to distinguish Lashley's hypothesis about action syntax from those linked to mirror neurons or embodiment. Lashley's model concerns a general and abstract notion of sequential hierarchy, and is only tied to motor actions insofar as these demand planning and control of a particular type. Lashley's model has nothing to do with grasping or recognizing graspable objects, and indeed involves structuring entities for which no such low-level action schemata are available. When applied to the role of the IFG, this leads to different empirical predictions from those of mirror-system hypotheses.

### Action syntax: beyond mirror neurons

Because localization is central to mirror-based arguments, precision regarding specific regions of the frontal cortex is key to the following discussion. The cytoarchitecture and connectivity of neighboring frontal regions changes drastically over small distances. These properties range from agranular cortex in motor and premotor regions (BA 4 and 6), to dysgranular BA 44, to granular BA 45 and anterior frontal cortex. The motor cortex generates strong descending motor pathways (e.g., the corticospinal tract), but BA 44/45 are both truly prefrontal, and clearly functionally distinct from motor regions. With broad or imprecise spatial localization, or multisubject averaging of activations, it is easy to confuse (or conflate) activations in one of these regions with that of the others.

Starting with motor and premotor activations, brain-imaging studies reveal clear and consistent activations for certain classes of words or images connoting either direct actions (“throw” or “kick”) or tools providing affordances for action

(“hammers,” “pliers”). Such motor-charged stimuli elicit (pre)motor activation; other more visually charged stimuli (e.g., animals) do not.<sup>80,91</sup> Recent studies<sup>35</sup> showing motor activation for particular components of language provide no special evidence for motor (or mirror neuron) involvement with language more broadly, because many linguistically structured concepts (color, clouds, beauty) have no motor component. Second, those few mirror neurons that have been directly documented in the human brain (during single-unit recording in 21 preoperative epilepsy patients) were found in the medial frontal (supplementary motor area) and the temporal (parahippocampal) cortex;<sup>93</sup> it is therefore possible that mirror neurons are much more widespread in human brains than in macaques. Simple observation of IFG activation in the human brain therefore does not specifically implicate mirror neurons, and more direct and precise tests are needed to distinguish between general frontal involvement and mirror neuron involvement.

A key testable prediction of mirror neuron-based hypotheses follows from the fact that, by definition, mirror neurons only exist for motor actions the subject can actually perform itself. To the extent that mirror neurons are required to understand or process some observed action, this predicts that unfamiliar or unproducible actions should not excite the mirror system, and should be uninterpretable (or at least difficult to process). This is rather difficult to test in the domain of language, because all normal humans are both producers and perceivers of language. Nonetheless, in brain-damaged patients with lesions in proposed sites of mirror neurons (the IFG and parietal regions), speech perception is not disturbed as predicted.<sup>94</sup> This study only found impaired speech perception when lesions involved auditory regions in the temporal lobe, inconsistent with motor theories of speech perception and of the need for mirror neurons to support accurate speech perception.

Additional data do not support the prediction that motor familiarity is required to elicit appropriate processing of speech stimuli. An infant optical imaging study<sup>95</sup> found frontal activation to speech-based patterns in newborns, despite their lack of experience in producing speech motor patterns themselves. Similarly, an artificial grammar-learning study by Bahlmann *et al.* employed abstract visual patterns, sequentially presented, which

had no motor (or verbal) component whatsoever. Items were arranged into either context-free patterns favoring hierarchy or finite-state patterns favoring pure sequence. The authors found activation of BA 44 only for the first case, supporting a role for this region in sequential hierarchy in the absence of motor-action understanding.<sup>40</sup>

Perhaps the most direct test of the mirroring prediction is possible with music, where many humans are devoted music listeners but lack the motor capability to play this music themselves. Mirror neuron enthusiasts often cite studies indicating motor involvement and Broca's activation during music perception by expert musicians on their own instruments.<sup>35</sup> But the crucial test is actually with nonmusicians, where the data are clear: Broca's activation is found during music perception, and musical syntax tasks in particular, even in nonmusicians, whose understanding of instrumental music presumably has little or no motor component.<sup>37,96</sup> Furthermore, although many findings are consistent with the general notion that music perception relies partially on motor and action-related circuitry,<sup>60,72,75</sup> most of the motor areas implicated are neither part of the putative mirror neuron system, nor specifically related to music production (e.g., singing or playing).

The dangers of rough localization and characterization are illustrated in an intriguing paper by Fazio *et al.*,<sup>54</sup> which examined the abilities of patients with frontal aphasia to correctly understand event sequences. Previous work<sup>81</sup> showed a particular deficit for prefrontally damaged patients for interpreting action. The Fazio study expanded this by first showing participants short video sequences of human actions or physical events and then presenting still images extracted from the videos on a touch screen. The patients' task was to properly sequence these images by touching them (requiring no overt language). The six patients in this study showed a general deficit relative to healthy controls, but were more impaired for sequencing human actions (e.g., a man grasping a bottle) than physical events (e.g., a bicycle falling down). Superimposing the MRI-determined lesion locations of patients upon one another revealed that the only region consistently damaged in all patients was centered on BA 44. However, all but one of these patients also had much more extensive damage, often stretching back into the parietal lobe. Thus, although intriguing,

these results do not support the strong inferences some draw about a specific role of Broca's area in this task,<sup>35</sup> which involves action sequencing but not hierarchical sequencing per se.

### Broca's area and hierarchical action planning

Much more convincing evidence for a specific role for Broca's area in hierarchical action, beyond music and language, comes from an elegant fMRI study by Koehlin and Jubault,<sup>97</sup> which introduced a task designed to discriminate simple sequencing from hierarchical sequencing. This work built upon the computationally explicit model of hierarchical action planning of Dehaene and Changeux,<sup>98</sup> which recognizes that at least three levels of nesting are needed to clearly discriminate simple gestures, motor sequences, and hierarchical plans. This model is based on the "Tower of London" task introduced by Shallice,<sup>42</sup> which involves moving colored beads on three posts from some starting state to an end state designated with a target image. This task can be smoothly varied in complexity, from an easy variant where the beads are simply moved directly to the goal, to challenging end states that require five or more moves, some of which require moving beads away from their final goal state.<sup>99</sup> Shallice found that patients with left prefrontal damage have difficulties specifically with these more complex problems,<sup>42</sup> and brain-imaging results are broadly consistent with prefrontal involvement in this task.<sup>100</sup> Dehaene's and Changeux's model specifies three levels of action: simple motor gestures (e.g., grasping a bead), operations (e.g., moving a bead from one post to another), and plans (nested sequences of operations to reach some specified goal state). Only the planning level provides clear evidence of hierarchical (nonchained) sequencing, particularly when it involves intermediate operations leading away from the goal. When the computational model was "lesioned" by removing planning units, these more difficult problems were specifically impaired, just as in Shallice's prefrontal patients.

Koehlin and Jubault<sup>96</sup> designed an analogous task. Repeated button presses (motor condition) or simple, overlearned left—right sequences are produced in the scanner, under different stimulus conditions. In the simple chunk condition, overlearned sequences alternated with the motor control sequencing task. But in the superordinate condition,

chunk sequences were arranged into higher level orders. All sequences were highly overlearned and automatized, so no learning occurred during scanning. Although simple motor acts elicited motor and premotor activity, higher order actions elicited more anterior activations within Broca's area, with transitions between chunks favoring BA 44, and termination of superordinate chunks activating BA 45. These data provide the clearest evidence to date of a specific role for Broca's area in planning hierarchically structured action.

We conclude that neuroimaging data support Lashley's hypothesis that music, language, and some types of action planning share a common substrate, and that Broca's area plays a key role in this distributed system. Although none of these data exclude a role for mirror neurons in action understanding in music or language, they argue against a necessity for motor knowledge to either understand sensory input or to elicit specific activity in the IFG. In contrast, these results remain consistent with Lashley's more abstract proposals, on the basis of computational similarities underlying planning of complex action in all three domains. The burning question thus remains concerning what, precisely, this supramodal system is doing at a computational level.

### Broca's area as Lashley's "scannable buffer"

Lashley emphasized that the core requirement of sequential hierarchy is that higher level processes and goals can be placed on hold during the execution of lower level processes. To accomplish such persistence, he suggested, certain neural circuits must act as a buffer scannable by other circuits. In computational terms, sequential hierarchies require the equivalent of a register (holding one element), a queue (multiple freely accessible elements) or a stack (holding multiple elements, but with restricted last-in-first-out access), and require that these forms of persistent memory be available to other ongoing computational processes. Psychologically, these are all different forms of working memory that provide temporary storage of intermediate results (analogous to a scratchpad or blackboard). Although a reflex chain of actions, where each action triggers the next, does not require such persistent intermediate storage, both processing and production of hierarchical sequences do. Furthermore, the more

complex and multileveled the hierarchy, the greater the capacity of the intermediate storage mechanism must be.

From a computational viewpoint, there is no doubt that effective, succinct computation requires such additional persistent storage. The most common means of achieving it in computer science is via a stack. For example, during the execution of complex computer programs, each call to a subroutine places a pointer to the current routine (and often intermediate results) onto the stack before branching to execute the subroutine. Once the subroutine completes its computation, typically returning some required result, execution of the original calling function can resume where it left off, by popping its pointer back off the stack. This can be embedded multiple times (with each level of nesting requiring additional stack space), and clearly the storage capacity of the stack limits the possible extent of embedding.

Both this storage capacity and restrictions on access influence the types of computations that are possible. Theoretical computer scientists have mapped out, in considerable detail, the computational limitations of registers versus stacks versus queues in abstract models of computation called *automata*. For instance, a finite-state automaton (FSA) is a system where discrete computational states are linked with no persistent memory. When such a system is augmented with additional working memory structures, new automata classes result. A push-down automaton (PDA), for example, is an FSA with an additional stack provided, whereas a Turing machine is an FSA with an endless queue (unrestricted order of access). Automata are models of processing, but important theorems link each of these types to specific classes of string sets ("languages") that can be generated or recognized by the corresponding type of automaton. Unsurprisingly, the less restricted the additional memory store of the automaton, the broader and less restricted the type of string sets it can handle. These automaton/language pairs can thus be arranged into a classification system, termed the formal language hierarchy or extended Chomsky hierarchy:<sup>101</sup> a fundamental theoretical concept in computer science.<sup>102–104</sup>

One appealing version of such stack-based computing would involve domain-specific templates and filler information, retrieved and maintained by a multidomain processor. For example the sentence "John likes apples" would involve retrieving

the template “X likes Y,” along with the specific elements “John” and “apples” (which might just as well be “Mary” and “chocolate”). The role of the scannable buffer here is simply to maintain the current template-element bindings during the processing of further perceptual information or motor planning. Such a mechanism could be used to deal with many types of templates, independent of format (e.g., verbal, melodic, and motor), and thus would be available for multiple domains.

Although the distinctions between registers, stacks, and queues are a major focus in computer science, from a psychological or neural point of view the relevance of such abstractions is less obvious. Lashley’s critical point was that hierarchical sequence processing requires some form of scannable intermediate storage, and limitations on this capacity obviously restrict the capacity of a system to handle hierarchical structures. Formally, the crucial distinction is between FSA-only systems (which can generate the so-called regular languages and have only a very limited and inflexible capacity to process hierarchy) and the augmented automata above these (e.g., PDAs or Turing machines), which correspond to the suprarregular languages.

A foundational result in computational linguistics is that natural languages require processing resources over and above those of an FSA, and are thus suprarregular.<sup>105,106</sup> Considerable debate has surrounded the question of how far beyond regular the natural languages go; current consensus is that languages go a bit beyond the capacities of a PDA and belong to the class termed *mildly context-sensitive grammars*,<sup>107,108</sup> requiring multiple stacks, or even a stack of stacks to be computed. With music, the situation is less clear, but some theorists suggest that music requires computational suprarregular resources including at least one stack<sup>109</sup>—the so-called context-free grammars.

Marrying these computational considerations with our prior conclusions on the basis of neural localization leads to the following hypothesis. Cortical resources in the IFG, including at least BA 44 and BA 45, implement a storage buffer scannable by other cortical and subcortical circuits subserving sequential behavior. This buffer is required to implement suprarregular hierarchical sequence processing, and its processing load increases with the depth and complexity of the hierarchy being processed.

This hypothesis, which has been entertained in related forms by multiple researchers,<sup>52,97,110</sup> extends Lashley’s speculations by positing a discrete neural locus implementing his scannable buffer and by specifying its suprarregular nature. Direct evidence that processing load increases with depth is provided by parametric studies,<sup>53,111,112</sup> and the localization aspect to the IFG is consistent with a wealth of neuroimaging data. This hypothesis does not require that the IFG is the only site of a scannable neural buffer (similar buffers may be implemented, for example, in hippocampus or basal ganglia circuitry). Nor does this hypothesis suggest that the stack or buffer represents all of the relevant data or computations: the idea is that it can activate and link representations generated in other brain regions (akin in computer terms to storing a pointer, rather than all of the data). Nonetheless, the involvement of Broca’s region in hierarchical sequencing in music and language, and in some complex action planning, seems abundantly clear. Koelsch has termed this idea the *syntactic equivalence hypothesis*.<sup>113</sup>

This hypothesis does not follow directly from the fact that cortical systems for both perception and action are laid out hierarchically. For if each level of a processing hierarchy mapped directly onto a neural hierarchy, implemented on the cortical surface, there would be strict limitations on the possible depth of hierarchical structures. For example, the data on hierarchical action planning above suggested a direct mapping from first-level motor gestures in posterior (motor and premotor) cortex and third-level superordinate planning to anterior (BA 45) regions. But prefrontal cortex cannot be extended forever, and a literal mapping would suggest a strict limit on the level of embedding possible in language, music, or action. Such a limit is not evident in action (even making coffee can be construed as involving four levels, and if interrupted by a phone call, five), nor in music or language. The well-known restriction of embedding to three levels in language, for example, applies only to a particular syntactic structure—center embedding—not to syntax or language as a whole.<sup>114,115</sup>

Having thus extended Lashley’s hypothesis to a specific neural circuit, we turn to Lashley’s final hypothesis: that the capacity to process musical and linguistic hierarchical sequences is unusually developed in humans, but builds upon and extends pre-existing action sequencing mechanisms present in

animals. Addressing this issue demands an examination of comparative data.

### The phylogenetic roots of hierarchical sequencing: comparative data

Many comparative studies have addressed animals' ability to produce and perceive hierarchical structure,<sup>116</sup> mostly in the auditory domain but also exploring visual capabilities.<sup>45,117</sup> Much of the recent animal work has adopted the computational principles outlined above, contrasting finite-state grammars with context-free grammars to see which (if any) animal species can go beyond the regular grammar level.

This comparative work built upon human experimental psychology research begun by George Miller in the late 1950s, in one of the first computerized testing laboratories at Harvard.<sup>118</sup> Rule-learning problems were generated and tested at different computational levels. The results led Miller to put forth the following hypothesis, which we call the *supraregular hypothesis*: adult humans have a proclivity to induce rule systems at context-free (or higher) levels, even when the data do not require such systems. That is, even stimulus sets that could be captured by regular grammars tend to be captured using supraregular grammars.

Miller's hypothesis suggests not only that humans are able to go beyond regular grammars, but also that human observers are biased to view sequences as hierarchically structured, even when the generating algorithm is a serial, finite-state system. Because all stimulus sets generated by finite-state grammars can also be parsed by context-free grammar, this may be a cognitive example of the aphorism "to a man with a hammer everything looks like a nail." We might think of this bias as a form of dendrophilia—an inordinate fondness for tree structures—in our species.<sup>119</sup>

Both hierarchy and supraregularity are clear and well defined mathematically. Unfortunately, it is not trivial to adjudicate between finite-state and context-free requirements on the basis of string processing alone.<sup>120</sup> Structure is an invisible (mental) construct, and can only be empirically probed indirectly (e.g., by observing acceptance and rejection of well-chosen string sets,<sup>121</sup> by adding prosodic cues at or between putative phrase boundaries,<sup>122</sup> or by observing reactions to clicks played at or between putative boundaries<sup>123</sup>). Such indirectness renders

empirical testing of these notions challenging, but by no means impossible.

Returning to animal research, existing studies paint a relatively consistent picture. Many species, from rats to songbirds to baboons, have well-developed sequencing capabilities that can be characterized as finite state systems. Such systems, for example, appear adequate to capture the syntactic structure seen in birdsong<sup>116</sup> or the recognition of simple long-distance dependency in squirrel monkeys.<sup>124</sup> In contrast, the capacity of animals to induce supraregular grammars remains contentious. Despite numerous examples of successful induction of regular grammars, only a few species have been claimed to go beyond this,<sup>116</sup> and both the data and the methods in these studies have been sharply questioned.<sup>125–127</sup> Perhaps the most convincing attempt to show supraregular abilities in a nonhuman species came from Gentner *et al.* in their operant conditioning work<sup>128</sup> with starlings (*Sturnus vulgaris*), songbirds with complex song and extended vocal learning in both females and males.<sup>129</sup> Gentner's *et al.* used a supraregular grammar denoted  $A^n B^n$ , which means that the number of A units must be precisely matched by the number of B units (in this case the A and B elements were two types of starling song motifs, termed *warbles* and *rattles*, presented acoustically). This grammar is provably supraregular, and allows multiple possible structural interpretations ranging from count-and-compare, to center-embedded, to cross-serial structure.<sup>121</sup> After very extensive training (between 9400 and 56,200 trials per bird, with a mean of 30,000 trials), three of four starlings tested with various probe strings showed evidence of having acquired the supraregular language  $A^n B^n$ . This suggested that, although quite difficult, it was at least possible for some birds to acquire a supraregular rule. However, a follow-up study in zebra finches<sup>125</sup> showed that such levels of performance could be explained by simpler rules and averaging together of the performance of different birds.

On the basis of these and other results (reviewed in Ref. 116), we can tentatively formulate a comparative extension of Miller's human-oriented supraregular hypothesis—the *supraregular exceptionality* hypothesis, which posits that, in contrast to humans, nonhuman animals display a particular difficulty inducing hierarchical patterns from strings of data, especially any structures requiring

computational resources above the finite-state level.

Evidence that an organism fails on suprarregular grammars (e.g.,  $A^nB^n$ , mirror grammars, or copy grammars<sup>101</sup>), although succeeding on regular grammars (e.g.,  $(AB)^n$ ,  $AB^*A$ , or sequential transition–probability grammars<sup>130,131</sup>) constitutes support for this hypothesis. Unlike Miller's human-specific suprarregular hypothesis, which has abundant support from human studies,<sup>40,118,132–134</sup> our suprarregular exceptionality hypothesis, which concerns nonhuman animals, remains tentative for several reasons. First, nearly every comparative study has used a different species, and only a handful of nonhuman species have been tested to date. This renders broad generalizations about animals impossible. Even statements about one species remain tentative: multiple empirical issues need to be considered in pattern-learning experiments that make it difficult for any single study to control for all interpretations. Finally, very few grammars have been tested (indeed, all animal-based claims for suprarregularity are based on a single grammar  $A^nB^n$ ), which renders generalizations to broad formal classes impossible. Although humans succeed on multiple suprarregular grammars (including copy and mirror grammars), to our knowledge no one has tested such grammars in nonhumans. Firm conclusions will require more research, on different species and different grammars, and replications across laboratories.

Nonetheless, these comparative data provide initial support for Lashley's notion that a hierarchical syntax of action is particularly well developed in our species. It seems too early to draw conclusions concerning the degree to which this proclivity builds upon preexisting suprarregular abilities in animals.<sup>135–137</sup> Certainly, we have no evidence of action syntax in animals even approaching the complexity of making a pot of coffee: the closest would be certain instances of tool use in some chimpanzee populations.<sup>138</sup> Before returning to the possible neural basis for expanded hierarchical abilities in humans, we need to briefly forestall misunderstanding by noting an area of pervasive confusion in the recent comparative literature.

### **Recursion and suprarregularity: an unfortunate confusion**

For unclear reasons much of the comparative literature just discussed conflates several distinct issues,

mixing up multiple computational abilities as if they were one. Our concern in this review is the processing of hierarchical sequences, defined as sequences with a tree-formed structure in which element order matters. Such structures are best processed by suprarregular systems. Unfortunately, much of the last decade's research based on formal language theory has conflated hierarchy and suprarregularity, in these precise senses, with various notions of recursion. *Recursion* refers to that broad class of computational processes in which a function calls itself (in computer science), or to those structures where the same hierarchical structure is repeated at multiple levels of the hierarchy ("self-embedding"). An example of a recursive structure is a fractal, where the same structure repeats itself ad infinitum at all hierarchical levels.<sup>139</sup>

Crucially, although all recursive trees are hierarchical, all hierarchies are not recursive. For many types of hierarchy discussed in this review, such as motor actions embedded within plans, it is unclear what self-embedding would even mean. Although a context-free grammar can be implemented recursively, for finite sets it need not be, and regular grammars can also be implemented recursively. The well-defined suprarregular divide, between finite-state automata and those suprarregular automata that have some additional scannable memory, is thus orthogonal to the issue of recursion. This issue has now been repeatedly discussed in the literature,<sup>121,140–144</sup> so we will not belabor it here. We can only hope that the unfortunate current tendency to substitute the ambiguous high-profile buzzword *recursive* for distinct (and more empirically tractable) concepts like *suprarregular* or *hierarchical* will eventually run its course.

### **Evolutionary expansion of Broca's region in humans**

We end with a brief review of recent comparative anatomical data pointing to a general expansion of prefrontal connectivity in our species, along with a more specific and pronounced increase in the size of Broca's area and changes in its pattern of connections.

The human brain is greatly expanded in size relative to other primates, including our nearest relatives the great apes (chimpanzees, gorillas, and orangutans).<sup>145</sup> However, ours are not the largest brains in the animal kingdom: the brains of

elephants and toothed whales like dolphins or orcas are larger.<sup>146</sup> Nor do human brains occupy the greatest relative proportion of total body weight: that distinction goes to small mammals like shrews.<sup>147,148</sup> It is only when brain size is considered relative to what would be predicted, given body size (the so-called *encephalization quotient*), that human brains reign supreme.<sup>149</sup> In these terms, although great apes have relatively large brains for mammals, humans are still quite exceptional, with a roughly threefold increase in total brain size relative to chimpanzees (or early fossil hominids like *Australopithecus*<sup>150</sup>).

Recent comparative neuroanatomical work allows us to go beyond these long-known facts to inquire whether human brain expansion affected specific brain regions and circuits. Although human frontal regions, like the cortex overall, are much larger than in other primates, whether this increase is disproportionate to overall size increase remains contentious. One of the first studies to use MRI in living primates to address this issue<sup>151</sup> compared the volume of the frontal lobe in four great ape species with humans, along with lesser ape (gibbon) and monkey (*Macaca* and *Cebus*) brains. They concluded that great apes in general have an expanded frontal cortex relative to gibbons or monkeys, but frontal lobe volume was not disproportionately different from humans relative to total brain volume. This is consistent with many findings that apes are cognitively superior to monkeys.<sup>152</sup> However, this study looked at the entire frontal lobe, as demarcated by the central sulcus (thus including both motor and anterior areas).

To examine the relative size of the prefrontal cortex specifically, Schoeneman *et al.* used MRIs of 11 primate species to examine the brain volume anterior to the corpus callosum.<sup>153</sup> A separate analysis of gray and white matter found a striking dissociation between these two components: human relative gray volumes differed significantly from only a few species, whereas white volumes were greater than all but two species. A subsequent analysis of gray:white size ratios indicates that any human expansion of the prefrontal cortex relative to total brain volume is almost entirely the result of an increase in white matter. This suggests that, in addition to their raw increase in absolute size, human prefrontal regions have become disproportionately connected to the rest of cortex. Whether human frontal lobes (gray

or white) are significantly and disproportionately larger than predicted depends on the comparison group (apes or monkeys). But differences from non-ape primates are established, whereas differences from apes remain controversial.<sup>154–157</sup> Furthermore, absolute frontal volume increase is undisputed, and relative white matter increases seem relatively clear.

To what extent do these increases in size and connectivity apply equally to all prefrontal regions? This question is challenging to answer, because it requires that multiple brains are analyzed using detailed cytoarchitectonic methods. In apes, such analysis has to date been performed only for prefrontal areas BA 10 (frontal pole), 44, and 45 (along with V1 and BA 13, part of the insula). Comparisons between the brains of 12 chimpanzees<sup>158</sup> with previous work in humans<sup>159</sup> revealed a striking finding: areas 44 and 45 are the most greatly expanded cortical areas yet identified in humans. Left areas 44 and 45 are six times larger in humans than in chimpanzees,<sup>158</sup> disproportionate to the roughly threefold increase in total brain size, or the 4.5-fold increase in frontal cortex in total. This important study shows that Broca's area in particular has expanded disproportionately since our divergence from chimpanzees (roughly 6 million years ago).

Regarding connectivity, a recent study exploited the technique of diffusion tensor imaging (DTI), which uses MRI to estimate, in intact brains, axonal connectivity between brain regions.<sup>160</sup> Applying DTI to human, macaque, and chimpanzee brains, Rilling *et al.*<sup>161</sup> found significant changes in the pattern of connectivity between Broca's area and posterior brain regions. Humans, and to some degree chimpanzees, showed connectivity between prefrontal and temporal regions via a dorsal pathway through the parietal cortex, but only in humans did this pathway (the arcuate fasciculus) have a strong uninterrupted connection to the posterior temporal cortex. In contrast, rhesus macaques showed a very weak dorsal pathway, whereas ventral connections between the temporal cortex and frontal regions via the insula dominated. These results again converge on the conclusion that it is not simply the size of prefrontal regions that has increased in humans; the pattern of connectivity has changed as well.<sup>162</sup>

### Conclusion: Broca meets Lashley

Combining the behaviorally anomalous status of human music and language relative to animal

communication systems with the comparative anatomical data and imaging data reviewed above, we begin to see the outlines of a neurally and biologically grounded hypothesis concerning the undeniable differences between humans and other animals that does justice to the very deep cognitive and neural foundations that we share with other vertebrates (including other primates). For each of Lashley's three questions, laid out in the introduction, current data support a tentative positive answer: processing hierarchical sequences appears to be a well-defined and neutrally localizable function, shared by musical and linguistic syntax (Q1), humans are unusually well developed in this ability (Q3), and it is at least plausible to suggest that this ability may inherit key components from some type of action syntax that predated the evolution of human music and language (Q2).

From a modern perspective, Lashley's action syntax can be localized to a set of widespread brain circuits that have, as a key hub, prefrontal regions centered on Broca's region. The prefrontal regions play a primitive role in the hierarchical planning and sequencing of action, a function presumably shared with other primates, and particularly with chimpanzees, whose consistent and intelligent use of tools has no peer among other nonhuman primates.<sup>163–165</sup> However, the apparent remit of this type of hierarchical planning increased greatly during human evolution, to include both perception and production of all types of hierarchical sequences. The most prominent additions to this Broca-centered action sequencing capacity were, by hypothesis, those two great human achievements: music and language.

From an evolutionary viewpoint, this idea provides for a certain continuity between humans and other primates (particularly chimpanzees) while acknowledging the drastic increases in certain cognitive capacities that have occurred since our divergence with chimpanzees. They also make sense of the fact that although prefrontal cortex and Broca's area are not novel brain regions, they have expanded disproportionately and have changed their patterns of connectivity to other brain regions during recent human evolution. In proposing this, we are fully aware that Broca's region is by no means the sole seat of language or music. Both capacities rely on a far-flung network of both cortical and subcortical brain regions (all of them, again, shared with other

primates). But Lashley's ideas bring into focus the precise cognitive and computational changes that might underlie our expanded capacities for planning, action, language, music, and thought itself, in the context of a wide array of shared capacities. This is, by hypothesis, our broadly developed ability and indeed proclivity to structure both action and perception hierarchically. The clear distinctions between currently popular mirror neuron-based hypotheses and the more abstract computational hypothesis of Lashley have the potential to drive more refined experimental procedures and to provide further empirical evidence relevant to understanding Broca's region and its function. We conclude that, although Lashley has been rarely cited in the large number of studies implicating Broca's region in action, music, and language, his ideas have withstood the passage of time well, and deserve renewed attention.

### Acknowledgments

We thank Daniel Bowling, Bruno Gingras, Stefan Koelsch, Marisa Hoeschele, and two anonymous reviewers for comments on a previous version of this manuscript. We also acknowledge support from ERC Advanced Grant #230604 "SOMACCA" and University of Vienna Research Cluster Grant "Shared Neural Resources" (to W.T.F.), and FCT Grant SFRH/BD/64206/2009 (M.M.).

### Conflicts of interest

The authors declare no conflicts of interest.

### References

1. Lashley, K. 1951. The problem of serial order in behavior. In *Cerebral Mechanisms in Behavior; the Hixon Symposium*. L.A. Jeffress, Ed.: 112–146. New York: Wiley.
2. Rosenbaum, D.A., et al. 2007. The problem of serial order in behavior: Lashley's legacy. *Hum. Mov. Sci.* **26**: 525–554.
3. Chomsky, N. 1959. A note on phrase structure grammars. *Inform. Control.* **2**: 393–395.
4. Chomsky, N. 1968. *Language and Mind*. New York: Harcourt, Brace & World.
5. Upshall, M. 1993. *Hutchinson Dictionary of Mathematics*. London: Brockhampton Press.
6. Illingworth, V. 1983. *Dictionary of Computing*. Oxford: Oxford University Press.
7. Altmann, C.E., H.H. Bühlhoff & Z. Kourtzi. 2003. Perceptual organization of local elements into global shapes in the human visual cortex. *Curr. Biol.* **13**: 342–349.
8. Aminoff, E., N. Gronau & M. Bar. 2006. The parahippocampal cortex mediates spatial and nonspatial associations. *Cereb. Cortex* **17**: 1493–1503.

9. Badre, D. 2008. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* **12**: 193–200.
10. Badre, D. & M. D'Esposito. 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* **10**: 659–669.
11. Badre, D., et al. 2009. Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nat. Neurosci.* **12**: 515–522.
12. Kikuchi, Y., B. Horwitz & M. Mishkin. 2010. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. *J. Neurosci.* **30**: 13021–13030.
13. Kourtzi, Z., et al. 2003. Integration of local features into global shapes: monkey and human fMRI studies. *Neuron* **37**: 333–346.
14. Kravitz, D.J., et al. 2013. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**: 26–49.
15. Krumbholz, K., et al. 2005. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. *Eur. J. Neurosci.* **21**: 230–238.
16. Mormann, F., et al. 2008. Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J. Neurosci.* **28**: 8865–8872.
17. Kravitz, D.J., et al. 2011. A new neural framework for visuo-spatial processing. *Nat. Rev. Neurosci.* **12**: 217–230.
18. Kumaran, D., H.L. Melo & E. Duzel. 2012. The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron* **76**: 653–666.
19. Schendan, H.E., et al. 2003. An fMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron* **37**: 1013–1025.
20. Meyer, P., et al. 2005. Language processing within the human medial temporal lobe. *Hippocampus*. **15**: 451–459.
21. Opitz, B. & A.D. Friederici. 2003. Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *Neuroimage*. **19**: 1730–1737.
22. Opitz, B. & A.D. Friederici. 2007. Neural basis of processing sequential and hierarchical syntactic structures. *Hum. Brain Mapping*. **28**: 585–592.
23. Eichenbaum, H., et al. 2012. Towards a functional organization of episodic memory in the medial temporal lobe. *Neurosci. Biobehav. Rev.* **36**: 1597–1608.
24. Ranganath, C. 2010. A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus* **20**: 1263–1290.
25. Hassabis, D., et al. 2007. Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U S A*. **104**: 1726–1731.
26. Baddeley, A. 2000. The episodic buffer: a new component of working memory? *Trends Cogn. Sci.* **4**: 417–423.
27. Henke, K. 2010. A model for memory systems based on processing modes rather than consciousness. *Nat. Rev. Neurosci.* **11**: 523–532.
28. Amunts, K., et al. 2010. Broca's region: novel organizational principles and multiple receptor mapping. *PLoS Biol.* **8**: 1000489.
29. Friederici, A.D. 2002. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* **6**: 78–84.
30. Friederici, A.D. 2011. The brain basis of language processing: from structure to function. *Physiol. Rev.* **91**: 1357–1392.
31. Petersson, K.M., C. Forkstam & M. Ingvar. 2004. Artificial syntactic violations activate Broca's region. *Cogn. Sci.* **28**: 383–407.
32. Forkstam, C., et al. 2006. Neural correlates of artificial syntactic structure classification. *Neuroimage* **32**: 956–967.
33. Koelsch, S., et al. 2013. Processing of hierarchical syntactic structure in music. *Proc. Natl. Acad. Sci. U S A*. **110**: 15443–15448.
34. Brown, S., M.J. Martinez & L.M. Parsons. 2006. Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur. J. Neurosci.* **23**: 2791–2803.
35. Fadiga, L., L. Craighero & A. D'Ausilio. 2009. Broca's area in language, action, and music. *Ann. New York Acad. Sci.* **1169**: 448–458.
36. Koelsch, S., B. Maess & A.D. Friederici. 2000. Musical syntax is processed in the area of Broca: an MEG study. *Neuroimage*. **11**: 56.
37. Maess, B., et al. 2001. Musical syntax is processed in Broca's area: an MEG study. *Nat. Neurosci.* **4**: 540–545.
38. Patel, A.D., et al. 2008. Musical syntactic processing in agrammatic Broca's aphasia. *Aphasiology* **22**: 776–789.
39. Sammler, D., S. Koelsch & A.D. Friederici. 2011. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex* **47**: 659–673.
40. Bahlmann, J., et al. 2009. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* **1298**: 161–170.
41. Luria, A.R. 1966. *Higher Cortical Functions in Man*. New York: Basic Books.
42. Shallice, T. 1982. Specific impairments of planning. *Philos. Trans. R. Soc. B.* **298**: 199–209.
43. Baddeley, A.D. 1986. *Working Memory*. Oxford: Clarendon Press.
44. Passingham, R.E. 1993. *The frontal lobes and voluntary action*. Oxford, UK: Oxford University Press.
45. Conway, C.M. & M.H. Christiansen. 2001. Sequential learning in non-human primates. *Trends Cogn. Sci.* **5**: 539–546.
46. Passingham, R.E. 1981. Broca's area and the origins of human vocal skill. *Philos. Trans. R. Soc. B.* **292**: 167–175.
47. Bookheimer, S. 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* **25**: 151–188.
48. Friederici, A.D. 2009. Pathways to language: fiber tracts in the human brain. *Trends Cogn. Sci.* **13**: 175–181.
49. Zurif, E.R., A. Caramazza & R. Myerson. 1972. Grammatical judgments of agrammatic aphasics. *Neuropsychologia* **10**(4): 405–417.
50. Hagoort, P. 2005. On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* **9**: 416–423.
51. Hickok, G. & D. Poeppel. 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**: 393–402.

52. Pulvermüller, F. 2010. Brain embodiment of syntax and grammar: discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain Lang.* **112**: 167–179.
53. Pallier, C., A.-D. Devauchelle & S. Dehaene. 2011. Cortical representation of the constituent structure of sentences. *Proc. Natl. Acad. Sci. U S A.* **108**: 2522–2527.
54. Fazio, P., et al. 2009. Encoding of human action in Broca's area. *Brain* **132**: 1980–1988.
55. Thompson-Schill, S.L. 2005. Dissecting the language organ: a new look at the role of Broca's area in language processing. In *Twenty-First Century Psycholinguistics: Four Cornerstones*. A. Cutler, Ed.: 173–190. London: Lawrence Erlbaum.
56. Thompson-Schill, S.L., et al. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci. U S A.* **94**: 14792–14797.
57. Patel, A.D. 2013. Sharing and nonsharing of brain resources for language and music. In *Language, Music, and the Brain: A Mysterious Relationship*. M.A. Arbib, Ed.: 329–355. Cambridge, Massachusetts: MIT Press.
58. Koelsch, S. 2013. Neural correlates of music perception. In *Language, Music, and the Brain: A Mysterious Relationship*, Vol. 10. M.A. Arbib, Ed.: 141–172. Cambridge, Massachusetts: MIT Press.
59. Janata, P. & L.M. Parsons. 2013. Neural mechanisms of music, singing, and dancing. In *Language, Music, and the Brain: A Mysterious Relationship*, Vol. 10. M.A. Arbib, Ed.: 307–328. Cambridge, Massachusetts: MIT Press.
60. Herholz, S.C., A.R. Halpern & R.J. Zatorre. 2012. Neuronal correlates of perception, imagery, and memory for familiar tunes. *J. Cogn. Neurosci.* **24**: 1382–1397.
61. Peretz, I. & M. Coltheart. 2003. Modularity of music processing. *Nat. Neurosci.* **6**: 688–691.
62. Patel, A.D. 1998. Syntactic processing in language and music: different cognitive operations, similar neural resources? *Musical Percept.* **16**: 27–42.
63. Koelsch, S., et al. 2005. Interaction between syntax processing in language and in music: an ERP study. *J. Cogn. Neurosci.* **17**: 1565–1577.
64. Slevc, L.R., J.C. Rosenberg & A.D. Patel. 2009. Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon. Bull. Rev.* **16**: 374–381.
65. Steinbeis, N. & S. Koelsch. 2008. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb. Cortex.* **18**: 1169–1178.
66. Fitch, W.T. 2006. The biology and evolution of music: a comparative perspective. *Cognition* **100**: 173–215.
67. Fitch, W.T. 2013. Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Front. Syst. Neurosci.* **7**: 1–16.
68. Abrams, D.A., et al. 2011. Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cereb. Cortex.* **21**: 1507–1518.
69. Rogalsky, C., et al. 2011. Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J. Neurosci.* **31**: 3843–3852.
70. Fedorenko, E., M.K. Behr & N. Kanwisher. 2011. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl. Acad. Sci. U S A.* **108**: 16428–16433.
71. Chen, J.L., R.J. Zatorre & V.B. Penhune. 2008. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J. Cogn. Neurosci.* **20**: 226–239.
72. Grahn, J.A. 2012. Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* **4**: 585–606.
73. Grahn, J.A. & M. Brett. 2007. Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* **19**: 893–906.
74. Wymbs, N.F. & S.T. Grafton. 2013. Contributions from the left PMd and the SMA during sequence retrieval as determined by depth of training. *Exp. Brain Res.* **224**: 49–58.
75. Janata, P. & S.T. Grafton. 2003. Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat. Neurosci.* **6**: 682–687.
76. Madison, G. 2006. Experiencing groove induced by music: consistency and phenomenology. *Music Percept.* **24**: 201–208.
77. Janata, P., S.T. Tomic & J.M. Haberman. 2012. Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol.: General.* **141**: 54–75.
78. Liberman, A.M. & I.G. Mattingly. 1985. The motor theory of speech perception revised. *Cognition* **21**: 1–36.
79. Diehl, R.L., A.J. Lotto & L.L. Holt. 2004. Speech perception. *Ann. Rev. Psychol.* **55**: 149–179.
80. Martin, A., et al. 1996. Neural correlates of category-specific knowledge. *Nature* **379**: 649–652.
81. Tranel, D., et al. 2003. Neural correlate of conceptual knowledge for actions. *Cogn. Neuropsychol.* **20**: 409–432.
82. Arbib, M.A. 2002. The mirror system, imitation, and the evolution of language. In *Imitation in Animals and Artifacts*. C. Nehaniv & K. Dautenhahn, Eds.: 229–280. Cambridge, MA: MIT Press.
83. Arbib, M.A. 2005. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**: 105–167.
84. Rizzolatti, G. & M.A. Arbib. 1998. Language within our grasp. *Trends Neurosci.* **21**: 188–194.
85. Condillac, É.B.D. 1971 (1747). *Essai sur l'origine des Connaissances Humaines*. Gainesville, FL: Scholar's Facsimiles and Reprints.
86. Corballis, M.C. 2002. *From Hand to Mouth: the Origins of Language*. Princeton: Princeton University Press.
87. Hewes, G.W. 1973. Primate communication and the gestural origin of language. *Curr. Anthropol.* **14**: 5–24.
88. Emmorey, K. 2005. Sign languages are problematic for a gestural origins theory of language evolution. *Behav. Brain Sci.* **28**: 130–131.
89. Kendon, A. 1991. Some considerations for a theory of language origins. *Man* **26**: 199–221.
90. MacNeilage, P.F. & B.L. Davis. 2005. The frame/content theory of evolution of speech: a comparison with a gestural-origins alternative. *Interact. Stud.* **6**: 173–199.

91. Toni, I., *et al.* 2008. Language beyond action. *J. Physiol.* **102**: 71–79.
92. Fitch, W.T. 2010. *The Evolution of Language*. Cambridge: Cambridge University Press.
93. Mukamel, R., *et al.* 2010. Single-neuron responses in humans during execution and observation of actions. *Curr. Biol.* **20**: 750–756.
94. Rogalsky, C., *et al.* 2011. Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase* **17**: 178–187.
95. Gervain, J., *et al.* 2008. The neonate brain detects speech structure. *Proc. Natl. Acad. Sci. U S A.* **105**: 14222–14227.
96. Koelsch, S., *et al.* 2002. Bach speaks: a cortical “language-network” serves the processing of music. *NeuroImage*. **17**: 956–966.
97. Koechlin, E. & T. Jubault. 2006. Broca’s area and the hierarchical organization of human behavior. *Neuron* **50**: 963–974.
98. Dehaene, S. & J.-P. Changeux. 1997. A hierarchical neuronal network for planning behavior. *Proc. Natl. Acad. Sci. U S A.* **94**: 13293–13298.
99. Berg, W.K., *et al.* 2010. Deconstructing the tower: parameters and predictors of problem difficulty on the Tower of London task. *Brain Cogn.* **72**: 472–482.
100. Newman, S.D., J.A. Greco & D. Lee. 2009. An fMRI study of the Tower of London: a look at problem structure differences. *Brain Res.* **1286**: 123–132.
101. Jäger, G. & J. Rogers. 2012. Formal language theory: refining the Chomsky Hierarchy. *Philos. Trans. R. Soc. B.* **267**: 1956–1970.
102. Gersting, J.L. 1999. *Mathematical Structures for Computer Science*. New York: W H Freeman.
103. Hopcroft, J.E., R. Motwani & J.D. Ullman. 2000. *Introduction to Automata Theory, Languages and Computation*. Reading, Massachusetts: Addison-Wesley.
104. Linz, P. 2001. *An Introduction to Formal Languages and Automata*. Sudbury, Massachusetts: Jones & Bartlett.
105. Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.
106. Levelt, W.J.M. 2008. *Formal Grammars in Linguistics and Psycholinguistics*. Amsterdam: John Benjamins.
107. Stabler, E.P. 2004. Varieties of crossing dependencies: structure dependence and mild context sensitivity. *Cogn. Sci.* **28**: 699–720.
108. Stabler, E.P. 2013. The epicenter of linguistic behavior. In *Language Down the Garden Path: The Cognitive and Biological Basis of Linguistic Structures*. M. Sanz, I. Laka & M.K. Tanenhaus, Eds.: 316–323. New York: Oxford University Press.
109. Rohrmeier, M. 2011. Towards a generative syntax of tonal harmony. *J. Math. Music.* **5**: 35–53.
110. Uddén, J. & J. Bahlmann. 2012. A rostral-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**: 2023–2032.
111. Braver, T.S., *et al.* 1997. A Parametric Study of Prefrontal Cortex Involvement in Human Working Memory. *NeuroImage*. **5**: 49–52.
112. Petersson, K.M., V. Folia & P. Hagoort. 2010. What artificial grammar learning reveals about the neurobiology of syntax. *Brain Lang.* **120**: 83–95.
113. Koelsch, S. 2012. *Brain and Music*. London, UK: John Wiley & Sons.
114. Miller, G.A. & N. Chomsky. 1963. Finitary models of language users. In *Handbook of Mathematical Psychology*. Vol. 35. R.D. Luce, R.R. Bush & E. Galanter, Eds.: 419–492. New York: John Wiley & Sons.
115. Bach, E., C. Brown & W. Marslen-Wilson. 1986. Crossed and nested dependencies in German and Dutch: a psycholinguistic study. *Lang. Cogn. Process.* **1**: 249–262.
116. ten Cate, C. & K. Okanoya. 2012. Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philos. Trans. R. Soc. B.* **367**: 1984–1994.
117. Stobbe, N., *et al.* 2012. Visual artificial grammar learning: comparative research on humans, kea (*Nestor notabilis*) and pigeons (*Columba livia*). *Philos. Trans. R. Soc. B.* **367**: 1995–2006.
118. Miller, G.A. 1967. Project Grammmarama. In *Psychology of Communication*. G.A. Miller, Ed. New York: Basic Books.
119. Fitch, W.T., A.D. Friederici & P. Hagoort. 2012. Pattern perception and computational complexity. *Philos. Trans. R. Soc. B.* **367**: 1925–1932.
120. Fitch, W.T. Toward a computational framework for cognitive biology: unifying approaches from cognitive neuroscience and comparative cognition. *Phys. Life Rev.* In press.
121. Fitch, W.T. & A.D. Friederici. 2012. Artificial grammar learning meets formal language theory: an overview. *Philos. Trans. R. Soc. B.* **367**: 1933–1955.
122. Morgan, J.L., R.P. Meier & E.L. Newport. 1987. Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cogn. Psychol.* **19**: 498–550.
123. Fodor, J.A., T.G. Bever & M.F. Garrett. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw-Hill.
124. Ravignani, A., *et al.* 2013. Action at a distance: dependency sensitivity in a New World primate. *Biol. Lett.* **9**: 20130852.
125. van Heijningen, C.A.A., *et al.* 2009. Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proc. Natl. Acad. Sci. U S A.* **106**: 20538–20543.
126. Beckers, G.J.L., *et al.* 2012. Birdsong neurolinguistics: songbird context-free grammar claim is premature. *NeuroReport*. **23**: 139–145.
127. Petkov, C.I. & E.D. Jarvis. 2012. Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* **4**: e12.
128. Gentner, T.Q., *et al.* 2006. Recursive syntactic pattern learning by songbirds. *Nature* **440**: 1204–1207.
129. Hausberger, M., *et al.* 1995. Song sharing reflects the social organization in a captive group of European starlings (*Sturnus vulgaris*). *J. Comp. Psychol.* **109**: 222–241.
130. Saffran, J., *et al.* 2008. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* **107**: 479–500.

131. Saffran, J.R., R.N. Aslin & E.L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* **274**: 1926–1928.
132. de Vries, M.H., *et al.* 2008. Syntactic structure and artificial grammar learning: the learnability of embedded hierarchical structures. *Cognition* **107**: 763–774.
133. Fitch, W.T. & M.D. Hauser. 2004. Computational constraints on syntactic processing in a nonhuman primate. *Science* **303**: 377–380.
134. Uddén, J., *et al.* 2012. Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: investigating the push-down stack model. *Cogn. Sci.* **2012**: 1–24.
135. Byrne, R.W. 2007. Clues to the origin of the human mind from primate observational field data. *Jpn. J. Anim. Psychol.* **57**: 1–14.
136. Hauser, M., N. Chomsky & W.T. Fitch. 2002. The language faculty: what is it, who has it, and how did it evolve? *Science* **298**: 1569–1579.
137. Johnson-Pynn, J., *et al.* 1999. Strategies used to combine seriated cups by chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*), and capuchins (*Cebus apella*). *J. Comp. Psychol.* **113**: 137–148.
138. Sanz, C., J. Call & D.B. Morgan. 2009. Design complexity in termite-fishing tools of chimpanzees. *Biol. Lett.* **5**: 293–296.
139. Mandelbrot, B.B. 1977. *The fractal geometry of nature*. New York: Freeman.
140. O'Donnell, T.J., M.D. Hauser & W.T. Fitch. 2005. Using mathematical models of language experimentally. *Trends Cogn. Sci.* **9**: 284–289.
141. Fitch, W.T. 2010. Three meanings of “recursion”: key distinctions for biolinguistics. In *The Evolution of Human Language: Biolinguistic Perspectives*. R. Larson, V. Déprez & H. Yamakido, Eds.: 73–90. Cambridge, UK: Cambridge University Press.
142. Fitch, W.T., M.D. Hauser & N. Chomsky. 2005. The evolution of the language faculty: clarifications and Implications. *Cognition*. **97**: 179–210.
143. Martins, M.D. 2012. Specific signatures of recursion. *Philos. Trans. R. Soc. B.* **367**: 2055–2064.
144. Van der Hulst, H. 2010. *Recursion and Human Language*. Berlin: De Gruyter/Mouton.
145. Huxley, T.H. 1863. *Man's Place in Nature*. Mineola, NY: Dover.
146. Jerison, H.J. 1961. Quantitative analysis of evolution of the brain in mammals. *Science* **133**: 1012–1014.
147. Stephan, H., H. Frahm & G. Baron. 1981. New and revised data on volumes of brain structures in insectivores and primates. *Folia. Primatol.* **35**: 1–29.
148. Deacon, T.W. 1990. Rethinking mammalian brain evolution. *Am. Zool.* **30**: 629–705.
149. Jerison, H.J. 1973. *Evolution of the Brain and Intelligence*. New York: Academic Press.
150. Holloway, R.L. 1996. Evolution of the human brain. In *Handbook of Human Symbolic Evolution*. A. Lock & C.R. Peters, Eds.: 74–108. Oxford: Clarendon Press.
151. Semendeferi, K., *et al.* 2002. Humans and great apes share a large frontal cortex. *Nat. Neurosci.* **5**: 272–276.
152. Deaner, R.O., C.P. van Schaik & V. Johnson. 2006. Do some taxa have better domain-general cognition than others? A meta-analysis of nonhuman primate studies. *Evol. Psychol.* **4**: 149–196.
153. Schoenemann, P.T., M.J. Sheehan & L.D. Glotzer. 2005. Prefrontal white matter volume is disproportionately larger in humans than in other primates. *Nat. Neurosci.* **8**: 242–253.
154. Barton, R.A. & C. Venditti. 2013. Human frontal lobes are not relatively large. *Proc. Natl. Acad. Sci. U S A.* **110**: 9000–9006.
155. Sakai, T., *et al.* 2011. Differential prefrontal white matter development in chimpanzees and humans. *Curr. Biol.* **21**: 1397–1402.
156. Sherwood, C.C., R.L. Holloway, K. Semendeferi & P.R. Hof. 2005. Is prefrontal white matter enlargement a human evolutionary specialization? *Nat. Neurosci.* **8**: 537–538.
157. Smaers, J.B., *et al.* 2010. Frontal white matter volume is associated with brain enlargement and higher structural connectivity in anthropoid primates. *PLoS One.* **5**: e9123.
158. Schenker, N.M., *et al.* 2010. Broca's area homologue in chimpanzees (*Pan troglodytes*): probabilistic mapping, asymmetry and comparison to humans. *Cereb. Cortex.* **20**: 730–742.
159. Amunts, K., *et al.* 1999. Broca's region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* **412**: 319–341.
160. Koch, M.A., D.G. Norris & M. Hund-Georgiadis. 2002. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *Neuroimage* **16**: 241–250.
161. Rilling, J.K., *et al.* 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* **11**: 426–428.
162. Catani, M. & M. Mesulam. 2008. The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* **44**: 953–961.
163. Goodall, J. 1986. *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, Massachusetts: Harvard University Press.
164. McGrew, W.C. & L.F. Marchant. 2001. Ethological studies of manual laterality in the chimpanzees of the Mahale Mountains, Tanzania. *Behaviour* **138**: 329–358.
165. Sugiyama, Y. & J. Koman. 1979. Tool-using and tool-making behavior in wild chimpanzees at Bossou, Guinea. *Primates* **20**: 513–524.