Review

Stone tools, language and the brain in human evolution

Dietrich Stout1,∗ and Thierry Chaminade2

1 Department of Anthropology, Emory University, 1557 Dickey Drive, Atlanta, GA 30322, USA
2 Mediterranean Institute of Cognitive Neuroscience, CNRS—Université Aix-Marseille 2, 29 Chemin Joseph Aiguier, 13420 Marseille Cedex 20, France

Long-standing speculations and more recent hypotheses propose a variety of possible evolutionary connections between language, gesture and tool use. These arguments have received important new support from neuroscientific research on praxis, observational action understanding and vocal language demonstrating substantial functional/anatomical overlap between these behaviours. However, valid reasons for scepticism remain as well as substantial differences in detail between alternative evolutionary hypotheses. Here, we review the current status of alternative ‘gestural’ and ‘technological’ hypotheses of language origins, drawing on current evidence of the neural bases of speech and tool use generally, and on recent studies of the neural correlates of Palaeolithic technology specifically.

Keywords: language evolution; mirror neuron; gesture; pedagogy; Oldowan; Acheulean

1. INTRODUCTION

Speculations regarding evolutionary relationships between toolmaking and language have a very long history. Darwin [1] himself observed that ‘To chip a flint into the rudest tool...demands the use of a perfect hand’ and that ‘the structure of the hand in this respect may be compared with that of the vocal organs’. This analogy was greatly extended by subsequent researchers, who described commonalities in the motor control of manipulation and articulation [2] and in the hierarchically structured serial ordering [3] of manual praxis and linguistic syntax [4–6]. Writing just a few years after Darwin, Engels [7] argued that language evolution was stimulated by ‘the development of the hand’, which led to increasing ‘mutual support and joint activity’ and finally gave ‘men in the making...something to say to each other’. This social thread was also picked up by subsequent workers, who considered the possible role of language in the transmission and coordination of early technologies [8–10], and suggested similarities between the sharing of arbitrary design concepts in the production of formal tool ‘types’ and the sharing of arbitrary symbolic associations in linguistic semantics [4,9].

In recent years, hypothetical links between vocal language and manual praxis have received new support from cognitive neuroscience. Although language processing was long viewed as a functionally specialized and anatomically discrete module within the brain, it is now clear that the so-called ‘language areas’ contribute to a wide array of non-linguistic behaviours [11], including tool use [12]. Indeed, one-to-one brain-behaviour mappings of complex functions like ‘language processing’ have largely been replaced by explanations of regional brain function in terms of more abstract computational properties [11] and context-specific interactions with anatomically distributed networks [13,14]. In this framework, it is expected that complex behaviours will map onto neural substrates in a flexible manner and that single regions will participate in multiple different functional networks [15,16]. From an evolutionary perspective, this presents an ideal context for the co-option of existing neural substrates to support new behavioural phenotypes (i.e. ‘exaptation’ [17]). The intersection of language and praxis networks in Broca’s area currently provides one of the best known examples of such complex functional overlap in human neocortex.

Broca’s area was originally identified as a discrete region of the left third frontal convolution specifically responsible for ‘the faculty of spoken language’ [18]. However, it is now recognized that frontal ‘language-relevant’ cortex extends across the entire inferior frontal gyrus (IFG) and contributes to a diverse range of linguistic functions involving the comprehension and production of syntactic, semantic and phonetic structure [19,20]. Furthermore, IFG is known to participate in a range of non-linguistic behaviours from object manipulation to sequence prediction, visual search, arithmetic and music [13,21,22]. It has been proposed that this superficial behavioural diversity stems from an underlying computational role of IFG in the supramodal processing of hierarchically structured information [23], leading to speculation that this function may have evolved first in the context of manual praxis before being co-opted to support other behaviours such as...
language [20]. Thus, current evidence and interpretation supports and refines various ‘technological hypotheses’ positing neural and evolutionary connections between language and technological praxis [2,4–6].

The fact that IFG participates in the perceptual comprehension as well as motor production of behaviour [24] has also attracted a great deal of attention. In monkeys, individual neurons in area F5, a putative Broca’s area homologue, have been shown to selectively respond to the performance of a grasping action and to the observation of a similar action performed by another individual [25]. It is widely believed that a homologous ‘motor resonance’ mechanism in humans enables understanding of the actions and intentions of others through a form of internal simulation [26]. This recalls earlier motor hypotheses of speech perception [27], and has been seen as an evolutionary precursor to the ability to make and recognize intentional communicative gestures [28].

The Mirror System Hypothesis (MSH) [29] proposes that this primitive action-matching system underwent successive evolutionary modifications to support imitation, pantomime, manual ‘protosign’ and ultimately vocal language, thus providing a neural underpinning for ‘gestural hypotheses’ [30] of language origins. The MSH does not specify the evolutionary pressures leading these adaptations, but the specific response of monkey F5 and human Broca’s area to hand–object interactions [31], the predominance of object manipulation and tool-use behaviours among putative (e.g. [32]) instances of primate cultural (i.e. imitative sensu lato) learning, and the importance of complementary gesture and speech in the human transfer of tool skills [33] are all directly compatible with earlier hypotheses identifying the transmission and coordination of tool use as a likely context for the evolution of intentional communication and language [7,9,10].

Despite this new supporting evidence, many unanswered questions and reasons for scepticism remain. As Holloway [4] cautioned long ago, any motor activity can be described as a hierarchically structured sequence of behavioural units. The hypothesis of a special evolutionary relationship between toolmaking and language predicts more particular overlap in information processing demands and/or neuroanatomical substrates between these two behaviours. Early optimism [9] notwithstanding, many Palaeolithic archaeologists have seen this as unlikely in the face of apparent cognitive dissimilarities between toolmaking and language. In particular, it has been argued [34–37] that toolmaking behaviour is not ‘syntactical’ in the linguistic sense because much of its structure derives from external physical constraints rather than internal rules, and that it is not ‘semantic’ in the linguistic sense because shared cultural conventions of tool manufacture are constrained by function and learned through imitation rather than being truly arbitrary and intentionally communicated in the way that shared symbolic reference is thought to be. However, others (e.g. [38,39]) have maintained that at least some Palaeolithic toolmaking methods are underdetermined by physical and functional constraints and that their cultural reproduction does imply sharing of abstract syntactical structures and semantic content.

The question of what exactly is shared during action observation and execution is also a key controversy in cognitive neuroscience, and one of particular relevance to the MSH. Although it has been argued that motor resonance is a sufficient mechanism for the sharing of intentions and the development of intersubjective understanding [40], others question its ability to convey this type of information [41] and particularly its relevance to intentional communication [42]. The MSH proposes a transitional ‘protosign’ stage of conventionalized, intentionally communicative pantomimes specifically to bridge this gap and establish the ‘semantic space’ necessary for vocal language to become adaptive [43]. Better understanding the kind of meaning communicated during the imitative ‘apprenticeship’ [6] learning of technological skills is thus of interest to archaeologists and cognitive scientists alike, and is critical to evaluating alternative hypotheses of language evolution.

In a recent series of articles, we have attempted to shed light on some of these unanswered questions, including: (i) the anatomical overlap of language and tool use in Broca’s area [12]; (ii) the neural correlates [44,45], manipulative complexity [46] and hierarchical organization [47] of specific Palaeolithic toolmaking methods; and (iii) the brain mechanisms involved in the observational understanding of these methods [48]. Here, we review these results and assess the current state of gestural and technological hypotheses of language origins.

2. CORTICAL NETWORKS FOR SPEECH AND TOOL USE

Speech and tool use are both goal-directed motor acts. Like other motor actions, their execution and comprehension rely on neural circuits integrating sensory perception and motor control (figure 1). An obvious difference between speech and tool use is that the former typically occurs in an auditory and vocal modality, whereas the latter is predominantly visuospatial, somato-sensory and manual. Nevertheless, there are important similarities in the way speech and tool-use networks are organized, including strong evidence of functional–anatomical overlap in IFG and, less decisively, in inferior parietal and posterior temporal cortex (PTC).

Evidence of such overlap is open to at least three alternative interpretations. First, it might be that the apparent functional overlap actually reflects the presence of distinct but closely adjacent fields resolvable only at a higher level of spatial resolution. In this case, function might still be rigidly fractionated in terms of modality, effector-system, cognitive process or some other organizing principle, but in a complexly distributed and interdigitated manner (e.g. [49]). Second, it might be that different overt behaviours do indeed use the same neural substrates, and that the underlying ‘function’ of the relevant cortex needs to be re-described in more abstract terms. Third, and perhaps most reasonably, it might be that relatively large fields of cortex can indeed be associated with particular abstract computational functions but that within these fields there will also be highly context-sensitive variation in the dynamic and overlapping neural groups [50] recruited by specific tasks. We follow Adolphs
extrastriate visual cortices to the posterior parietal lobe supports visuospatial–motor transformations for action, whereas a ‘ventral stream’ from occipital to ventral and lateral temporal cortices is involved in mapping visual percepts to stored semantic knowledge about tool function and use. The confluence of these streams in the posterior [57] and/or anterior [59,60] inferior parietal cortex is thought to provide the integration of action and semantic knowledge required for the skilful use of familiar tools. This information is communicated to the premotor cortices of the frontal lobe, which are classically (e.g. [61]) seen as responsible for generating sequential action plans to be sent to primary motor cortex for execution. However, it is increasingly apparent that information flow within these frontal-posterior action circuits is bi-directional, with frontal ‘motor’ areas influencing perception of action [62] and posterior ‘sensory’ areas involved in coding specific motor acts [63]. Within this sensorimotor continuum, IFG appears to play a critical role assembling action elements into hierarchically structured sequences during motor production [64] and perceptual comprehension [65] of goal-oriented actions, especially those involving objects [66].

It has recently [67] been proposed that speech displays a similar two-stream organization. In this model, a dorsal stream flowing from the superior temporal auditory cortex to a vocal tract auditory-motor integration area at the parietal–temporal junction and on to posterior parts of Broca’s area support sensorimotor transformations for articulation. A ventral stream from superior to PTC and on to more anterior parts of Broca’s area maps auditory percepts to stored semantic representations. Much as in tool use, it is thought that this sensorimotor and semantic information is integrated in bi-directional frontal-posterior action circuits [20] linking parietal and temporal cortex to IFG [68,69], with IFG acting as a ‘unification space’ [19] for the assembly of lexical and phonetic elements into hierarchically structured sequences during speech production and language comprehension.

The most clear-cut distinctions between speech and tool use lie at the level of primary sensory and motor cortices, as expected for behaviours relying on different sensory modalities and somatic effectors. Intermediate processing stages display more similarities, including a closely analogous bi-directional frontal-posterior architecture in which sensorimotor and semantic elements are integrated and assembled into meaningful, goal-directed action sequences. For example, inferior parietal cortex in particular seems to play a common role in generating sensorimotor transformations for both speech and tool-use networks.

(b) Inferior parietal lobe
It has been proposed that parietal function may be anatomically fractionated into parallel effector systems [70,71]. For example, cortex in the vicinity of the parietal-temporal lobe junction (ventral supramarginal gyrus/posterior planum temporale) has recently been characterized as a sensorimotor integration area for the vocal tract [72,73], whereas sensorimotor integration for manual prehension has long been associated with more anterior portions of inferior

(a) Two ‘two-stream’ accounts
Tool use is currently understood within the framework of a ‘two-streams’ account of visual perception [56–58]. A ‘dorsal stream’ flowing from occipital
parietal cortex and intraparietal sulcus [74]. Thus, parietal speech and tool-use regions might perform similar computational functions but remain distinguishable owing to reliance on different effector systems. This interpretation is consistent with evidence that producing a melody manually (using a piano) rather than vocally (by humming) results in a shift of activation from the parietal-temporal junction to the anterior intraparietal sulcus [72] and that phonetic processing of a visuospatial/manual (sign) language produces anterior inferior parietal activation comparable with that involved in pantomimes of object use [75].

On the other hand, there is a substantial literature linking lesions in the vicinity of the parietal-temporal junction (posterior supramarginal gyrus and angular gyrus) to ideomotor apraxia [76], a disorder of skilled manual action that includes tool use [77]. Imaging studies similarly report activations of posterior inferior parietal cortex in response to viewing and naming tools [78], imagining the prehension of graspable objects [79], imitating object manipulation [80] and planning everyday tool use [81]. Conversely, anterior inferior parietal cortex has been associated with tasks involving (vocal) phonological short-term memory [82] and discrimination [83]. Such evidence suggests that tool-relevant and language-relevant cortex are quite widespread and co-extensive in the inferior parietal lobe and supports a general characterization of the inferior parietal lobe as a supramodal processing region involved in diverse auditory-motor [72,73]), tactile-motor [84,85] and visual-motor [79,81] transformations.

One framework that can help make sense of this supramodal processing is the computational model for motor control relying on internal models. Briefly, internal models are neural mechanisms that represent relationships between motor command and their sensory consequences. Forward models predict the sensory consequences of an executed movement, and can be used to cancel the perception of the sensory consequences of our own actions, and are paired to inverse models that map the desired sensory consequences (the goal) to the motor commands that can efficiently lead to these consequences [86]. The inferior parietal cortex has repeatedly been associated with such integration of sensory and motor information, for example, in the central cancellation of the sensory consequences of self-tickling in the parietal operculum [84], and the ventral supramarginal gyrus’ involvement in object manipulation [87] and subvocal articulation for speech perception [73].

Such integration is also critical to imitation, in which the sensory consequences of the others’ actions must be matched to appropriate motor commands for self-execution [88], and numerous studies have confirmed inferior parietal cortex involvement in imitation (e.g. [89,90]). Inferior parietal cortex appears to be especially important for the imitation of skilled actions with objects [90], perhaps reflecting a specific role in representing the body schema [89] in relation to the complex prehensile and functional properties of hand-held tools [90]. Inferior parietal cortex is similarly involved in vocal imitation [91], and lesions of this region are associated with conduction aphasia leading to deficits in speech repetition and production [92]. This suggests not only that inferior parietal cortex plays an analogous role integrating perception and action for both tool use and speech, but also that this contribution may be important for imitative processes involved in the social transmission of both technology and language. In any case, current evidence certainly does not suggest that the distinction between ‘linguistic’ and ‘technological’ tasks is a natural break-point for fractionating inferior parietal function. To the contrary, the motor control aspect of both tasks and consequent similarities in their underlying computational architecture provide an integrated explanation for inferior parietal involvement in the domains of language and manipulation.

### (c) Posterior temporal lobe

Another region of possible functional/anatomical overlap is the PTC (figure 1a). Generally speaking, PTC is involved in mapping diverse sensory percepts to supramodal semantic representations, for example in the association of speech sounds with lexical information [67,73,93,94] or the association of visually presented tools with functional movement patterns [95–97]. Paralleling the broader dorsal/ventral ‘stream’ distinction discussed above, PTC displays a rough functional gradient from superior regions representing biological motion to inferior regions representing object form.

Thus, the superior temporal gyrus/superior temporal sulcus responds to sensory consequences of biological movements, including the auditory consequences of discrete speech gestures [98] and the visible patterning of intentional face, hand and body motions [99,100]; the cortex spanning the superior temporal sulcus/ middle temporal gyrus supports the crossmodal integration of object form and motion cues [100,101]; and the inferior temporal gyrus is involved in the supramodal representation of object form [102] independent of motion [103].

These supramodal representations are ‘semantic’ in the sense that they constitute general knowledge of objects and motions that is not constrained to specific instances or exemplars [104] and can be recruited for tasks ranging from linguistic reference, to picture recognition and action performance [20]. Indeed, it is increasingly apparent that linguistic reference is supported by category-specific semantic circuits involving many of the same brain regions involved in non-linguistic perception and action [11,20]. It is thus unsurprising that some of the best evidence of neural overlap between language and tool use comes from the semantic processing of tool words [105]. This overlap occurs especially in posterior middle temporal gyrus, a region commonly activated by tool-related tasks [106], and may be easily understood in terms of the distributed, category-specific organization of semantic memory generally, rather than any special relationship between language and tool use.

Interestingly, however, posterior middle temporal gyrus is also one of several areas commonly activated during auditory sentence comprehension, especially when deciding if sentences are semantically plausible [73,94]. Sentences used in such studies have not been explicitly controlled for the presence/absence of
manipulable objects, but are certainly not limited to instances of tool use (e.g. ‘the moon ripens the tree’s branches’ [107]). This suggests that posterior middle temporal gyrus function may be of more general relevance to the semantic processing of language. For example, one hypothesis posits a pre-linguistic origin for sentential predicate-argument structure out of a more general semantic system for the representation of objects, actions and properties [108]. Along these lines, a recent study [109] reported direct overlap between visually presented ‘symbolic gestures’ (e.g. downward motion with open hands) and their spoken English glosses (‘settle down’) in posterior middle temporal gyrus, providing additional support for a characterization of this region as part of a more generalized semiotic system.

Many questions remain about the specific functional/anatomical organization of the brain’s semantic systems [110,111] but, as in sensorimotor processing in inferior parietal cortex, there is little evidence that the distinction between ‘linguistic’ and ‘technological’ content/processes is a natural one for fractionating posterior temporal function. Posterior middle temporal gyrus in particular stands out as a focal point of overlap between tool use and linguistic reference, perhaps reflecting shared neural mechanisms and evolutionary history [108,109].

**d) Inferior frontal gyrus**

Perhaps, the best documented overlap between speech and tool use occurs in IFG. This includes evidence of direct overlap between verb production and the observation of object-directed actions [112] and between tool-use action execution (using pencils, scissors and chopsticks) and language comprehension (story listening) [12]. This overlap is consistent with the now widely held view that IFG acts as a supramodal processor for hierarchically structured sequential information (e.g. [21]), characterized by a posterior–anterior processing gradient of increasing abstraction [23,113,114]. This gradient, running from the ventral premotor cortex of the precentral gyrus/sulcus through the IFG *pars opercularis* to *pars triangularis*, is evident both structurally and functionally. Anatomically, the increasing representation of an internal granular layer from the agranular motor cortex through the dysgranular premotor cortex to the granular prefrontal cortex of the IFG reflects an increase in local, recurrent connections thought to be important for the processing of incoming information [115]. This is complemented by analyses of IFG connectivity using diffusion tensor imaging [68,116,117] and resting-state activity correlation [69], which confirm the more narrow sensorimotor profile of ventral premotor cortex and show the greater connectivity of more anterior IFG with supramodal regions of posterior parietal and temporal cortex (see §2b,c) via the arcuate fasciculus.

Functionally, a wide variety of experimental manipulations [23,113,114] provide evidence of a gradient from relatively concrete stimulus-response mapping in posterior IFG to increasingly abstract context-sensitive action selection and association with conceptual/semantic information in mid-to-anterior IFG. It has been suggested that this supramodal gradient tracks the localization of phonological, syntactic and semantic language processing [19,69], as well as increasingly abstract representations of manual action [118]. Such a parallel organization is illustrated by numerous studies, for example in reports that ventral premotor activation is associated both with the kinematics of basic hand–object interactions [119] and with phonological processing [120], *pars opercularis* with simple tool-use action sequences [12] and linguistic syntax [121] and *pars triangularis* with more complex actions [122] and syntactic/semantic integration [123]. Across modalities, IFG activation increases with the complexity of tasks/stimuli presented at a particular level of abstraction, for example in the increased activation of *pars opercularis* in response to more syntactically complex sentences [121] and to the observation of more motorically complex manual actions [124]. There is thus good evidence for a supramodal fractionation of function in IFG but, as in the inferior parietal and PTC, clear distinctions between language- and tool-relevant networks are not readily apparent. Indeed, evidence of direct functional overlap [12] provides strong support for the hypothesis that these networks are, at least in part, coextensive.

**e) Lateralization of function**

Although both language and tool use have classically been associated with left-dominant networks [11,81,106], there is increasing awareness of the important and distinctive contributions of the right hemisphere. In the case of linguistic processing, there is evidence of right hemisphere dominance for affective prosody and context-dependent meaning (i.e. discourse level processing) [11,125,126], while in the case of tool use, the right hemisphere appears to play a key role in coordinating protracted, multi-step, manual action sequences [127,128]. In both cases, right hemisphere contributions pertain to the larger scale spatio-temporal and/or conceptual integration of behaviour, which may help to explain why these contributions have been less apparent in neuroscientific and neuropsychological investigations focusing on smaller scale (e.g. phonological, lexico-semantic, syntactic) language processing or on the simple use of everyday tools (e.g. pantomiming the use of a hammer or comb).

In keeping with this general characterization of hemispheric difference, damage to right inferior parietal lobe is commonly associated with large-scale spatial neglect, whereas left inferior parietal damage produces ideomotor apraxia, a disorder of discrete action execution. Importantly, deficits following right inferior parietal lesions are not limited to spatial neglect of the contralateral visual field but include non-lateralized impairments of spatial working memory as well as selective and sustained attention on both spatial and non-spatial tasks, including auditory as well as visual stimuli [129]. This suggests a more general, cross-modal role for the right inferior parietal lobe in the integration of perception and action over time, and is consistent with evidence of right inferior parietal involvement in processing affective prosody [126,130], imitating speech rate during repetition [91], imitating...
the sequential order of manual actions [128,131] and representing action outcomes independent of behavioural means [132].

An analogous pattern of functional lateralization is apparent in the temporal lobe. For example, a recent meta-analysis [133] highlighted right posterior temporal lobe involvement in context-dependent semantic integration, contrasting this with left hemisphere dominance for more discrete lexico-semantic tasks (e.g. object naming). This is consistent with an earlier proposal that ‘coarse coding’ of semantic information in the right hemisphere (i.e. stimuli generate a large number of weak associations) facilitates the identification of distant semantic relations during discourse comprehension, whereas left hemisphere ‘fine coding’ (fewer, stronger associations) facilitates rapid and constrained response selection. In the visuomotor modality, right PTC is implicated in the perception of biological motion [134] and consequent attribution of intentions [135], inferential processes that rely on the identification of complex, spatio-temporally extended patterns of relative motion. In contrast, left PTC is preferentially responsive to the simpler, rigid motions of tools [95] and appears to support the binding of synchronous perceptual attributes into discrete, cross-modal object representations [136].

Finally, although left IFG dominance for phonological and syntactical processing is well-known, IFG involvement in hierarchical behaviour organisation is clearly bilateral [23]. Right IFG is more specifically linked with the contextual processing of linguistic semantics [125] and affective prosody [137] and with task-set switching (i.e. updating action plans) in response to the perception of contextually relevant stimuli [138,139]. This is again consistent with the suggestion that there is a general difference in hemispheric-processing styles, with the left being specialized for rapid, small-scale action control and the right for large-scale, longer duration integrative functions [15,140,141]. Indeed, this hemispheric ‘division of labour’ may be reflected anatomically in the greater global interconnectedness of the right hemisphere when compared with the more discrete, nodal organization of the left hemisphere [142]. This structural asymmetry appears to be shared with macaques [142], in keeping with the hypothesis that hemispheric specialization predates both language and tool use [143]; however, a recently reported rightward asymmetry of pathways connecting posterior inferior parietal cortex to frontal premotor cortex may reflect more specific human adaptations for toolmaking [58].

3. STONE TOOLMAKING AND BRAIN EVOLUTION

The similarity of cognitive processes and cortical networks involved in speech and tool use suggests that these behaviours are best seen as special cases in the more general domain of complex, goal-oriented action. This is exactly what would be predicted by hypotheses that posit specific co-evolutionary relationships between language and tool use (e.g. [4,6]), but does not distinguish them from gestural origin hypotheses stipulating a central role for explicitly communicative, rather than simply praxic, action [29]. At issue is the behavioural context of uniquely human evolutionary developments that occurred since the last common ancestor with chimpanzees and which are thus largely inaccessible to comparative analysis. To resolve such questions, it is necessary to turn to the more direct evidence of human behavioural evolution offered by the archaeological record.

Palaeolithic stone tools provide a relatively abundant and continuous record of behavioural change over the past 2.5 Myr that is of direct relevance to technological hypotheses of language origins. Reconstruction of the necessary behaviours involved in the production and use of particular tool types can provide evidence for the emergence of cognitive processes, like those reviewed above, that are also important for language. This in turn requires an interpretive framework for deriving implied cognitive capacities from observed technological behaviours (e.g. [144,145]). We have attempted to develop such a framework by identifying the neural correlates of particular Palaeolithic toolmaking activities using [18F]-fluorodeoxyglucose positron emission tomography (FDG-PET) to assess brain activation during actual tool production [44,45] and functional magnetic resonance imaging (fMRI) to identify activation during the observation of toolmaking action [48].

We focused on two technologies, ‘Oldowan’ and ‘Late Acheulean’, that bracket the beginning and end of the Lower Palaeolithic, encompassing the first approximately 2.2 Myr (90%) of the archaeological record. Oldowan toolmaking is the earliest (2.6 Myr old [146]) known human technology and is accomplished by striking sharp stone ‘flakes’ from a cobbles ‘core’ held in the non-dominant (hereafter left) hand through direct percussion with a ‘hammerstone’ held in the right hand. Late Acheulean toolmaking is a much more complicated method appearing about 700 000 years ago and involving, among other things, the intentional shaping of cores into thin and symmetrical teardrop-shaped tools called ‘handaxes’ [47]. We compared these technologies: (i) with a simple bimanual percussive control task in order to identify any distinctive demands associated with the controlled fracture of stone, and (ii) with each other in order to identify neural correlates of the increasing technological complexity documented by the archaeological record.

(a) Oldowan toolmaking

Results (figure 2) indicate that Oldowan toolmaking is especially demanding of ‘dorsal stream’ structures (§2a) involved in visuomotor grasp coordination, including anterior inferior parietal lobe and ventral premotor cortex but not more anterior IFG [44]. This is consistent both with behavioural evidence of the sensorimotor [147,148] and manipulative [46] complexity of Oldowan knapping, and with the concrete simplicity [149–151] and limited hierarchical depth [47] of Oldowan action sequences. Attempts to train a modern bonobo to make Oldowan tools [152] similarly indicate a relatively easy comprehension of the overall action plan but continuing difficulties with ‘lower-level’ perceptual-motor coordination and affordance detection. In sum, the appearance of Oldowan tools in the archaeological
controlled fracture to remove large, thin flakes that travel more than half-way across the tool surface without also removing large portions of the tool edge. This is facilitated by preparation of the striking surface through small-scale chipping and/or abrasion before percussion, creating a long-range functional dependency between temporally and structurally discrete operations. At more abstract/superordinate levels of organization, Late Acheulean toolmaking may also involve functional dependencies between consecutive flake removals and between different technological ‘sub-goals’ (e.g. edging, thinning, shaping) creating further long-range dependencies and ‘syntactical’ complexity.

Unfortunately, the study by Makuuchi et al. [153] did not examine right hemisphere activity and so it is not known whether portions of right dorsal pars triangularis activated by Late Acheulean toolmaking participate similarly in language-relevant working memory processes. As reviewed above (§2e), right IFG is known to be preferentially involved in larger scale discourse and affective language processing as well as in switching between different task sets in response to contextually relevant perceptual cues. Furthermore, right IFG may be preferentially involved with visuospatial as opposed to phonological working memory [154,155]. Preferential activation of right IFG during Late Acheulean toolmaking, a complex visuospatial task involving perceptually driven shifts between distinct task sets associated with particular sub-goals, appears likely to reflect these distinctive right hemisphere-processing characteristics. Further support for this interpretation comes from a recent study [46] that used a data glove to record digital joint angles in the left hand during experimental Palaeolithic toolmaking. Results showed that, although toolmaking in general is manipulatively complex, Late Acheulean left-hand manipulation is no more complex than that already present in the Oldowan. This indicates that increased right IFG involvement in Late Acheulean toolmaking does not arise from increased manipulative complexity in the contralateral hand and must instead be explained in terms of the higher order behavioural and cognitive control characteristics of the right hemisphere.

The archaeologically attested ability of Late Acheulean hominins to implement hierarchically complex, multi-stage action sequences during handaxe production thus provides evidence of cognitive control processes that are computationally and anatomically similar to some of those involved in modern human discourse-level language processing. This provides a second behaviourally and chronologically grounded functional/anatomical link between technological and linguistic capacities, further extending the plausible context for co-evolutionary interactions (e.g. behavioural, developmental and/or evolutionary co-option). Notably, this link is independent of putative resonance mechanisms and communicative intentions and thus additional to rather than exclusive of gestural hypotheses.

4. INTENTIONAL COMMUNICATION

Experimental studies of Lower Palaeolithic tool production reviewed in §3 establish plausible evolutionary
links between individual technological praxis and particular aspects of speech and language processing. They do not, however, directly address the origins of intentional, referential communication that are the real focus of gestural hypotheses. The MSH in particular proposes that a ‘protosign’ system of intentionally communicative manual gestures, itself derived through the conventionalization of iconic pantomimes, provided a necessary scaffold for the later emergence of (proto-)speech. Technological pedagogy does represent one particularly likely context for the deployment of such pantomimes and protosigns [6] but this is not stipulated by the MSH. An alternative hypothesis [39] is that technological pedagogy in itself, including intentional demonstration and ostensive gestures [156] but not pantomime or conventionalized protosign, would have been an adequate scaffold for the evolution of intentional vocal communication. The MSH maintains that pantomime is fundamentally different from praxis because pantomime requires the observer to infer action goals and thus can be used to intentionally influence the thoughts of another individual (i.e. to communicate information). Praxis is considered insufficient for this purpose because it remains directly tied to observable instrumental goals, thus making pantomime a necessary transitional stage in the evolutionary sequence. The alternative ‘technological pedagogy’ hypothesis proposes that in sufficiently complex praxis, goals are so distal and abstract that they must be inferred rather than observed. This provides a context for purposeful communication through demonstrations intended to impart generalizable (i.e. semantic) knowledge about technological means and goals [156], without necessarily involving pantomime. Thus, the technological pedagogy hypothesis removes a major theoretical motivation for positing a transitional pantomime stage but is not itself incompatible with the presence of such a stage.

A key prediction of the technological pedagogy hypothesis is that observation of complex technological praxis, without accompanying linguistic or pantomimic contextualization, should be sufficient to induce high-level goal inference. It is not obvious that this should be the case, because the very ‘opacity’ and ambiguity of the goals involved raises questions about the extent to which they can be shared through simple observation. It has been proposed that motor resonance is a sufficient foundation for such sharing [40], but this is open to question [42]. To investigate this issue in the specific context of Lower Palaeolithic technological transmission, we collected fMRI data from subjects of varying expertise observing an expert demonstrator producing Oldowan and Late Acheulean tools [47]. At the first level of analysis, contrasts with a simple percussive control condition produced activations remarkably similar to those observed in FDG-PET studies of toolmaking action execution [44,45], including the association of right anterior IFG activation with Acheulean but not Oldowan toolmaking. This corroborates previous results and confirms the general importance of resonance mechanisms in toolmaking observation. In subsequent analyses, we found that technologically naive subjects responded to relatively low-level action elements in the stimuli (involving posterior IFG) consistent with the MSH account of praxic action observation. However, we also found that expert subjects, specifically when viewing the more teleologically complex Late Acheulean action sequences, activated portions of rostral anterior medial prefrontal cortex (figure 3) associated with the attribution of intentions [157]. These effects of expertise and technological complexity suggest a model of complex action understanding in which the iterative refinement of internal models through alternating observation (i.e. inverse aspect of internal models) and behavioural approximation (i.e. practice comparing forward models with real feedback) allows for the construction of shared pragmatic skills and teleological understanding. The specific association of Late Acheulean action observation with inference of higher level intentions provides support for the technological pedagogy hypothesis and links it with a specific, archaeologically visible context.

5. CONCLUSION
Accumulating evidence is increasingly supportive of technological hypotheses of language origins, and goes a long way towards allaying concerns that the similarity in the hierarchical, combinatorial organization of the two domains is a superficial one or that the ‘initiative’ learning of toolmaking skills is fundamentally distinct from intentional communication. In particular, evidence of intention attribution during the observation of stone toolmaking provides support for a ‘technological pedagogy’ hypothesis, which proposes that intentional pedagogical demonstration could have provided an adequate scaffold for the evolution of intentional vocal communication. This hypothesis is consistent with the widespread view that increasing reliance on social learning and pedagogy was a key factor in hominin brain and cognitive evolution [158–160] and removes one of the major motivations for positing a transitional pantomime stage as seen in current formulations of the MSH. Importantly, however, the technological pedagogy hypothesis is not incompatible with the presence of such a stage.
Interestingly, functional imaging studies of Lower Palaeolithic toolmaking have yet to reveal significant activation of ‘ventral stream’ semantic representations in the posterior temporal lobes. This may be because experimental paradigms to date have strongly emphasized the ‘dorsal stream’ visuo-motor action aspects of tool production. However, if this trend continues in more diverse experimental manipulations, it may provide some support for the view that Lower Palaeolithic technology is relatively lacking in semantic content [35,36], and suggest that this aspect of modern human cognition evolved later and/or in a different behavioural context.

We thank James Steele for organizing and editing this volume as well as Ralph Holloway and an anonymous reviewer for helpful comments. fMRI and data glove research discussed here was funded by the European Union Project HANDTOMOUTH.

REFERENCES

34 Graves, P. 1994 Flakes and ladders: what the archaeological record cannot tell us about the origins of language.


