In this Issue

From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language

Papers of a Theme Issue compiled and edited by James Steele, Pier Francesco Ferrari and Leonardo Fogassi

The world's first science journal

12 January 2012

Publishing Editor: Claire Rawlinson

(tel: +44 (0)20 7451 2602; fax: +44 (0)20 7976 1837;

philtransb@royalsociety.org)

Production Editor: Radha Dutia

6–9 Carlton House Terrace, London SW1Y 5AG, UK

rstb.royalsocietypublishing.org

AIMS AND SCOPE

Each issue of *Phil. Trans. R. Soc. B* is devoted to a specific area of the biological sciences. This area will define a research frontier that is advancing rapidly, often bridging traditional disciplines. *Phil. Trans. R. Soc. B* is essential reading for scientists working across the biological sciences. In particular, the journal is focused on the following four cluster areas: **neuroscience and cognition; organismal and evolutionary biology; cell and developmental biology; and health and disease**. As well as theme issues, the journal publishes papers from the Royal Society's biological discussion meetings. For information on submitting a proposal for a theme issue, consult the journal's website at **rstb.royalsocietypublishing.org**.

SUBSCRIPTIONS

In 2012 *Phil. Trans. R. Soc. B* (ISSN 0962-8436) will be published twice a month. Full details of subscriptions and single issue sales may be obtained either by contacting our journal fulfilment agent, Portland Customer Services, Commerce Way, Colchester CO2 8HP; tel: +44 (0)1206 796351; fax: +44 (0)1206 799331; email: sales@portland-services.com or by visiting our website at http://royalsocietypublishing.org/info/subscriptions. The Royal Society is a Registered Charity No. 207043.

Subscription prices 2012 calendar year	Europe	USA & Canada	All other countries
Electronic access only	£2284/€2968	US$4320	£2466/US$4421
Printed version plus electronic access	£2969/€3858	US$5615	£3205/US$5747

From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language

Papers of a Theme Issue compiled and edited by James Steele, Pier Francesco Ferrari and Leonardo Fogassi

Contents

Editorial

Current developments at *Philosophical Transactions of the Royal Society B*

This has been a very good year for the journal. The impact factor has increased from 5.1 to 6.05, and the journal is now fourth in the ISI category of 'Biology'. Of course impact factor is only one indication of the role or the value of a journal, but it is here an encouraging indication that, despite the increase in capacity to 24 issues each year, we are continuing to attract excellent guest editors and authors.

At the annual meeting of the editorial board, we conducted a SWOT analysis, to consider what aspect of the journal might need attention in the near future, and we have also discussed the general perceptions of the journal with a number of scientists. As a result, two areas on which we are focusing to broaden the reach of the journal, are the balance between different subject areas and the geographical spread of the contributors. Issues focusing on topics related to biochemistry and molecular, cell and developmental biology are showing a steady increase in frequency. This brings this area into more equal balance with the other three areas on which the journal focuses, with a corresponding slight decrease in the representation of organismal biology. Partly by appointments to the Editorial Board, we also aim to bring the geographical provenance of contributors more into line with the global distribution of research activity.

Much of the strength of the journal comes from its publication of the Royal Society Discussion Meetings. These issues are invariably focused on topics of great current interest, and often with great relevance to public policy. Recent examples include our plastic age [1] and new vaccines for global health [2]. We have also been successful in eliciting excellent proposals for the four annual rounds of editorial commissioning, and our current acceptance rate for these stands at 42 per cent. These issues have also been of great interest and have ranged from food security [3] to mammal conservation strategies [4] to military medicine [5]. One article on habitat preference [6] was even awarded the 2011 'Wildlife Publications Award—Outstanding Article'.

Finally, it is a great pleasure to highlight the generous and expert support that the journal receives from inside and outside the Royal Society. The members of the Editorial Board are a constant source of expertise, energy and ideas, and their input is vital in the process of commissioning and reviewing each issue. Joanna Bolesworth has done an excellent job in her temporary position as commissioning editor, and we all wish her well in her new position, and are also delighted to welcome back Claire Rawlinson from maternity leave. We look forward to another successful year for the journal.

Linda Partridge*   *2011*

University College London, Darwin Building, Gower Street, London WC1E 6BT UK
*l.partridge@ucl.ac.uk

REFERENCES

1 Thompson, R. C., Swan, S. H., Moore, C. J. & vom Saal, F. S. 2009 Our plastic age. *Phil. Trans. R. Soc. B* **364**, 1973–1976. (doi:10.1098/rstb.2009.0054)

2 Jenkins, D. G. & Ricklefs, R. E. 2011 Biogeography and ecology: two views of one world. *Phil. Trans. R. Soc. B* **366**, 2331–2335. (doi:10.1098/rstb.2011.0064)

3 Godfray, H. C. J. *et al.* 2010 The future of the global food system. *Phil. Trans. R. Soc. B* **365**, 2769–2777. (doi:10.1098/rstb.2010.0180)

4 Rondinini, C., Rodrigues, A. S. L. & Boitani, L. 2011 The key elements of a comprehensive global mammal conservation strategy. *Phil. Trans. R. Soc. B* **366**, 2591–2597. (doi:10.1098/rstb.2011.0111)

5 Evans, G. & Lillywhite, L. 2011 Introduction and overview. *Phil. Trans. R. Soc. B* **366**, 124–126. (doi:10.1098/rstb.2010.0245)

6 Beyer, H. L., Haydon, D. T., Morales, J. M., Frair, J. L., Hebblewhite, M., Mitchell, M. & Matthiopoulos, J. 2010 The interpretation of habitat preference metrics under use-availability designs. *Phil. Trans. R. Soc. B* **365**, 2245–2254. (doi:10.1098/rstb.2010.0083)

Introduction

From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language

James Steele[1,*], Pier Francesco Ferrari[2,3] and Leonardo Fogassi[3,4]

[1]*AHRC Centre for the Evolution of Cultural Diversity, Institute of Archaeology, University College London, 31-34 Gordon Square, London WC1H 0PY, UK*
[2]*Dipartimento di Biologia Evolutiva e Funzionale, Università di Parma, via Usberti 11/a, 43124 Parma, Italy*
[3]*Dipartimento di Neuroscienze and Istituto Italiano di Tecnologia (IIT), Brain Center for Motor and Social Cognition, via Volturno 39, 43125 Parma, Italy*
[4]*Dipartimento di Psicologia, Università di Parma, Borgo Carissimi 10, 43121 Parma, Italy*

The papers in this Special Issue examine tool use and manual gestures in primates as a window on the evolution of the human capacity for language. Neurophysiological research has supported the hypothesis of a close association between some aspects of human action organization and of language representation, in both phonology and semantics. Tool use provides an excellent experimental context to investigate analogies between action organization and linguistic syntax. Contributors report and contextualize experimental evidence from monkeys, great apes, humans and fossil hominins, and consider the nature and the extent of overlaps between the neural representations of tool use, manual gestures and linguistic processes.

Keywords: action organization; macaque; chimpanzee; mirror neurons; human evolution; palaeolithic

1. HISTORICAL BACKGROUND TO THE PRESENT ISSUE

It has been recognized since at least Darwin's day that the human hand may have evolved adaptively to facilitate the control of tools [1, p.138], and that the human vocal tract has evolved to facilitate the articulatory gestures of spoken language [1, p.138]. But how closely coupled were these adaptive trends in the hominin lineage? Scientists have frequently considered the possibility of common underlying organizing principles in the neurophysiology of (usually spoken) language and of manual praxis, focusing, for example, on the domain-general implications of primate encephalization [2], on parallel schedules of development across domains during human ontogeny [3], on similarities in hemispheric lateralization of function [4], or on the emergence of gestural communication as an evolutionary precursor of speech [5–8]. However, in earlier formulations, arguments for such a coupling were often complicated by clinical observations of dissociations between deficits in the linguistic and praxic domains, as well as by cases of divergent functional lateralization in healthy subjects; while to many linguists, the analogy between linguistic syntax and action organization has sometimes seemed too loosely defined to carry much interpretive weight.

The papers in this Special Issue examine tool use and manual gestures in primates as a window on the evolution of the human capacity for language. Two quite recent scientific developments make this an opportune moment to revisit this topic almost 20 years after Gibson and Ingold edited the ground-breaking *Tools, Language and Cognition in Human Evolution* (1993), which addressed a similar theme. The first is the now-widespread clinical and experimental use of methods, such as functional magnetic resonance imaging (fMRI), that were still in their infancy when earlier reviews such as Gibson & Ingold's [9] appeared. Such methods enable highly targeted hypothesis testing in both clinical and non-clinical settings, and can very usefully complement evidence obtained with longer established techniques. The second is the development of a novel, coherent and experimentally well-supported neurophysiological hypothesis of a common architecture for processing certain key aspects of manual actions and of language, namely the 'mirror neuron system' or 'mirror system' hypothesis [6,8,10,11]. Discovery of the mirror neuron system has shed considerable light on the functional properties of a fronto-parietal network of predominantly motor-related

* Author for correspondence (j.steele@ucl.ac.uk).

brain regions involved in action organization. Localized in their initial discovery to area F5 of the monkey's premotor cortex [12,13], a probable homologue of Broca's area in humans, and then found also in the inferior parietal cortex [14,15], mirror neurons fire during both action execution and action observation. This capacity has been considered the basis for action understanding. The presence of a parieto-frontal mirror system has then been demonstrated in humans [16–19]. In particular, it has been shown that this system is strongly activated during imitation [20] and it has been suggested, on the basis of several studies, that it plays an important role in speech comprehension [21].

Historically, paleoneurological work on endocasts of fossil hominins has emphasized the expansion of cortical language areas, notably Broca's area, as a distinctive structural marker associated with the emergence of human language capacities in earlier hominins (*Homo habilis* [22]; *Homo ergaster* [23]; but cf. [24,25]). However, a recent tracer study of the Broca's area homologue in living non-human primates indicates similarities to humans in connectivity and network architecture that may have provided early hominins with pre-adaptations for language [26]. Functionally, mirror neurons in area F5 in monkeys are activated by manual grasping actions as well as by ingestive and communicative orofacial gestures [27,28]; observations in human subjects have meanwhile also shown that Broca's area contains motor representations of hand movements as well as of speech-related actions (cf. [29]). This evidence has suggested to some scientists that human speech and language could have evolved by co-opting neurophysiological mechanisms involved in the organization of manipulative and ingestive actions. Subsequent work has supported the hypothesis of a close association in humans between some aspects of action and of language representation, in both phonology and semantics [6,21,27,30]. Nevertheless, a significant theoretical problem remains for any hypothesis that would derive language evolutionarily from action organization: namely, whether or not the action system can provide a sufficiently close analogue to linguistic syntax [11,31,32].

The parts of the Special Issue related to manual gestures focus primarily on manual actions involving tools. Tool use provides an excellent experimental context in which to investigate the analogy with linguistic syntax, for several reasons. Tools extend the effector organ (the hand and arm), and in complex tool use (defined by Johnson-Frey [33] as tool use that 'converts the movements of the hands into qualitatively different mechanical actions'), tools provide a greater range of possible operations than can be achieved with the innate reaching and grasping capability of the hand alone. This requires both semantic knowledge of individual tools' functions, and a generative set of rules for their effective use. Complex tool use typically also requires asymmetrical coordinated bimanual action (in which each hand plays a complementary role; [34]), which has been found to be the most reliable elicitor of population-level right-handedness in captive African apes [35]. Asymmetrical bimanual coordinated actions provide a context for hierarchical embedding, with the discrete but complementary actions of each hand

needing to be described in a nested action syntax; while long sequences of such actions organized towards a larger goal also create long-range dependencies (where a preparatory action at one time step is meaningful only in relation with another action that is executed at a later time step). Finally, there is an extensive archaeological record of hominin tool manufacture and use, which can be examined in tandem with the fossil anatomical evidence of the evolving hominin brain, hands and vocal tract to assess theories of the coupled or decoupled evolutionary history of our human capacities in the two domains.

2. CURRENT RESEARCH THEMES AS ILLUSTRATED BY CONTRIBUTORS TO THIS ISSUE

Among the most significant recent discoveries concerning the neurophysiology of action organization in primates are the discoveries of mirror neurons (as noted above; cf. [12]), and of the learned incorporation of the tool into an extended representation of the effector organ in the body schema [36]. Both discoveries were made in macaque monkeys. Two papers in this Special Issue build on these discoveries, and extrapolate evolutionary and comparative insights from observations of monkeys' tool-use learning. Iriki & Taoka [37] propose that the abstract cognitive functions of the inferior parietal cortex in humans derive from an expansion of areas originally involved in computing sensorimotor transformations for reaching and grasping actions, and emphasize the evolutionary importance of cortical plasticity (and the learned incorporation of the tool into the body schema), as seen in the learning-induced changes in the cortical micro-architecture of monkeys trained in tool-using tasks. They develop a speculative hypothesis for the evolution of increasing cognitive abstraction in tool use, and suggest that the brain mechanisms that subserve tool use, located in the parietal cortex, may bridge the gap between gesture and language by exploiting the same principles of spatial information processing to realize novel mental functions that are detached from body constraints.

In a more specific experimental context, Macellini *et al.* [38] meanwhile demonstrate the ability of macaques both to learn functional tool properties and then generalize them to novel objects, and to generalize functional tool use to novel tasks. However, when investigating the possibility of tool-use learning by observation of a demonstrator, Macellini *et al.* also find that macaques do not appear to be able to translate the visual presentation of a novel tool-using action demonstrated by an experienced third party into the production of a corresponding motor action themselves, although some forms of facilitation of tool interaction are present. As a speculation, they conclude that the common sequential organization of tool actions and speech and the overlap of activation, for both functions, of ventral premotor cortex and Broca's area, suggest that a basic organization of the motor system for hand and mouth actions has been exploited for the emergence of new functions that rely on the same mechanisms.

Two papers report experimental evidence of action organization in tool use by captive chimpanzees that

may also have some relevance to the evolution of language. It is sometimes suggested that there may be common mechanisms involved in vocal tract gestural units and in manual action units at the level of motor control. Calvin [39] suggested that aimed throwing of stone projectiles by hominins could have provided a pre-adaptation for speech motor control, because of the demands this action makes for precision in movement timing; Calvin also noted that skilful hammering (as in nut-cracking) requires similar patterns of brachiomanual coordination. In a case of social tool use (captive apes throwing faeces or wet chow at human visitors as they pass by the enclosure, with the projectile acting as a tool to elicit a desired reaction from the visitors), Hopkins *et al.* [40] report findings of associations in chimpanzees between aimed throwing ability and communicative ability. More interestingly, they examined whether specific brain structures could somehow relate to such behavioural skills. The findings have been quite surprising. In fact, they found a correlation between aimed throwing ability and white-to-grey matter ratios in the homologue of Broca's area and in the motor-hand area of the precentral gyrus (with the effects more pronounced in the hemisphere contralateral to the preferred throwing hand). The same workers have also found that in captive chimpanzees, aimed throwing ability is associated with greater size of the posterior cerebellum [41].

No study has yet been conducted on the extent of any brain morphological correlates of chimpanzee individual ability with a non-social tool in a nut-cracking task, which has been described as 'probably the most demanding manipulatory technique yet known to be performed by wild chimpanzees' [42, p.174]. However, Frey & Povinelli [43] found evidence that chimpanzees display anticipatory grip selection in a task involving a sequence of acts aimed to extract a piece of food using a tool (a dowel). Critical for the task was the type of grip used to grasp the object, as it revealed the capacity to anticipate the forthcoming task. In humans, this anticipatory ability is linked to activation of a network implicated in response selection including frontal and parietal regions as well as the bilateral cerebellum (which is likely to be involved in feed-forward predictions of the sensory consequences and motor costs of a motor action). They briefly note the possible analogy with the phenomenon of coarticulation in gestural phonology.

Stone tools provide the longest and best-preserved archaeological record of the evolution of tool use in hominins, and there have been numerous attempts to discern indirect evidence of the emergence of language in the stone tool record ([4,44,45]; but cf. [46]). However, a necessary first step is to gain a clearer understanding of the organization of actions that would have been required to produce and use the tools that archaeologists recover. Two papers report the use of experimental archaeological techniques (the replication of Paleolithic stone tools) to elucidate contrasts and similarities between stone tool use in different tasks. To assess what was distinctive about skilled tool use by early stone tool-making humans (when compared with present-day chimpanzees), Bril *et al.* [47] compare and contrast features of action production and task

organization in nut-cracking by chimpanzee and human subjects and in conchoidal fracturing of stone by human subjects, finding that the stone knapping task (replicating early archaeologically attested Oldowan techniques of flake removal from a single-platform core) is much more complex. They suggest that understanding human brain evolution as it affected skilled action execution in the stone-flaking task requires us to focus not on particular cortical areas in isolation, but rather on the coordinated evolution of different components of cortico-cerebellar systems. In particular, the marked expansion of the frontal cortico-cerebellar system in chimpanzees and humans appears to be consistent with their increased social learning capacities, exemplified in their similar learning strategies of fine motor skills such as tool use.

Stout & Chaminade [48] meanwhile use functional brain imaging to contrast cortical aspects of action organization in a similar Oldowan stone-flaking task with that involved in production of a later Lower Palaeolithic tool type, the Late Acheulean handaxe. Whereas the Oldowan task activates cortical areas involved in visuo-motor grasp coordination (including anterior inferior parietal lobe and ventral premotor cortex), but not the inferior frontal gyrus (IFG), the Late Acheulean task also activates the dorsal right IFG (*pars triangularis*), an area associated with more abstract action representation and greater hierarchical task complexity and with possible involvement of lateralized visuospatial working memory. This seems to reflect the relatively complex goal hierarchy of the Late Acheulean task, which involves both a greater number of discrete sequential knapping events, and long-range dependencies between individual events in an extended sequence. Stout & Chaminade review also 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 actions based on Palaeolithic technology.

The brain activation patterns of human subjects replicating the stone tool technology of Neanderthals have not yet been studied experimentally. However, Ambrose [49] notes the appearance by about 300 000 yrs BP of composite tools such as spears with hafted stone points made by Neanderthals and suggests that their assembly rules may be analogous to linguistic grammars. In addition, Neanderthals also appear to have been predominantly right-handed [4], suggesting the presence of a human-like left cerebral lateralization of function. It is, therefore, interesting to ask whether or not this species was also capable of human-like speech. Barney *et al.* [50] attempt to estimate the potential of the Neanderthal vocal tract to produce human-like articulatory gestures, concluding that the principal contrast between this species and modern humans lies in the more pronounced facial flattening of modern human skull morphology and the associated reduced length of the front (oral) resonating cavity. They make some progress in the difficult task of reconstructing this extinct species' vocal tracts, although their results do not resolve the question of whether or not this contrast with modern human facial architecture would have compromised Neanderthals' speech potential.

As mentioned above, Stout & Chaminade explicitly draw an analogy between action organization in a complex task such as production of a Lower Palaeolithic stone handaxe, and linguistic syntax. At a more general level, Pastra & Aloimonos [51] develop a framework for analysing the grammar of naturally occurring actions, suggesting that the key features of a minimalist generative grammar (nested and tail recursion, and 'merge' and 'move' operations) must also characterize the generative grammar of action, particularly for actions involving tools (where a set of unimanual-sequential and bimanual-synchronous movements may be necessary to set up a framework for final action execution to achieve the goal). The examples they give of nested recursion all also involve asymmetric coordinated bimanual actions (e.g. 'grasp with hand$_1$ knife, pin with knife bread—grasp with hand$_2$ fork, pin with fork cheese, lick with tongue cheese—bite with teeth bread', where hand$_1$ and hand$_2$ are the two hands, and the underlined sequence is (Pastra & Aloimonos propose) an example of a recursively nested action structure. Pastra & Aloimonos's action grammar may bring us closer to understanding the commonalities between action organization in tool-using tasks, and linguistic syntax (for additional recent discussions from alternative perspectives see Glenberg & Gallese [52] and Tettamanti & Moro [53]).

Communicative manual gestures have often been invoked as an evolutionary bridge between instrumental actions and syntactically ordered human vocal communication [5,7], although Stout & Chaminade [48] suggest that parsing of complex manual tool-use sequences during social imitation might have provided such a bridge for earlier hominins without the need to invoke a separate communicative gestural stage. Liebal & Call [54] note that the difficulty of categorically differentiating actions from communicative manual gestures in great apes may relate to the fact that many gestures are derived from non-communicative actions through phylogenetic or ontogenetic ritualization. Social learning represents a third mechanism whereby gestures can emerge out of actions. Liebal & Call suggest that given such a continuum between action and gesture, the heuristic classification of movements as communicative gestures requires the presence of some or all of the following features: motoric ineffectiveness, waiting for a response, gaze alternation and persistence. They cite a recent summary of a systematic comparison of gestures in apes and macaques by Call & Tomasello [55], which found that chimpanzees and orangutans more often incorporate objects in their gestures. As Liebal & Call note, this correlates with these species' greater propensity to use tools in the wild, and may therefore be indicative of a common neural substrate for tool use and gestural communication. In terms of continuities with human language, they also note some evidence that chimpanzees show population-level right handedness for manual gesturing; however, great ape communicative gestures are still typically imperative, dyadic and lacking in abstraction. In contrast, Cartmill *et al.* [56] discuss the uniqueness of human representational gestures—gestures that often resemble the actions on objects which they represent, but which are not in themselves motorically effective. They suggest that whereas in non-human primates gestures are typically abbreviated versions of actions that lack symbolic abstraction, humans have the ability both to deploy more abstract representational gestures (influenced by individual experience) and to use these in support of cognitive problem-solving, as well as in social communication. Cartmill *et al.* note, for example, that familiar instrumental actions can become routinized to simulate and represent problem-solving strategies in those task domains, and suggest that focusing exclusively on the communicative function of gestures obscures their role in support of the gesturer's own cognition.

Finally, Roby-Brami *et al.* [57], in an evolutionarily oriented review of aphasic and apraxic syndromes in humans, point out that transitive gestures (those involving a tool or other object) are more complex than intransitive gestures (which do not, and which are typically communicative, such as waving goodbye), and that deficits in transitive gestures are also more closely associated with classical apraxic syndromes. They observe more specifically that deficits in pantomiming of tool-using actions are associated with damage to the left IFG, while deficits in intransitive gesture (although less well understood) appear to be less closely linked to impairments of the left cerebral hemisphere. Roby-Brami *et al.* also note that cerebral lateralization for praxis is more strongly linked to language dominance than to manual preference, indicating commonalities at the level of semantics and conceptual knowledge, and that there is evidence for a convergence between language and praxis at the syntactic level. Although some clinical evidence of double dissociation in the incidence of aphasia and apraxia suggest that the two left-hemisphere-lateralized systems are functionally distinct, there are clearly substantial overlaps between the neural networks subserving language and praxis.

The papers in this Special Issue demonstrate a wide range of approaches to the study of primate tool use and to the action–language relationship. Five of the papers report work carried out as part of the HandTo-Mouth project, which was funded by the European Commission's Sixth Framework NEST Pathfinder scheme (cf. papers by Macellini *et al.*, Bril *et al.*, Stout & Chaminade, Barney *et al.* and Roby-Brami *et al.*). Six additional papers were contributed by scientists working outside the framework of that particular project, but whose work overlapped with and complemented it. Co-authors of three of the latter papers (Frey, Iriki and Pastra) had also acted as advisers or external reviewers at earlier stages of the HandToMouth project. Undoubtedly, there remain many unsolved problems, and there are numerous additional research dimensions (such as the search for precursors in living non-human primates of human cortical control specifically of vocal gestures) that could not be explored here. Nevertheless, we believe that these papers collectively make a coherent and substantial contribution to our understanding of the evolution of tool use and language, and we sincerely thank all their authors for their support and participation.

REFERENCES

1 Darwin, C. 1871 *The descent of man, and selection in relation to sex.* London, UK: John Murray.
2 Gibson, K. R. 2002 Evolution of human intelligence: the roles of brain size and mental construction. *Brain Behav. Evol.* **59**, 10–20. (doi:10.1159/000063730)
3 Greenfield, P. M. 1991 Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**, 531–551. (doi:10.1017/S0140525X00071235)
4 Steele, J. & Uomini, N. 2009 Can the archaeology of manual specialization tell us anything about language evolution? A survey of the state of play. *Camb. Archaeol. J.* **19**, 97–110. (doi:10.1017/S0959774309000067)
5 Corballis, M. C. 2003 From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* **26**, 199–260.
6 Fogassi, L. & Ferrari, P. F. 2007 Mirror neurons and the evolution of embodied language. *Curr. Dir. Psychol. Sci.* **16**, 136–141. (doi:10.1111/j.1467-8721.2007.00491.x)
7 Hewes, G. W. 1973 Primate communication and the gestural origin of language. *Curr. Anthropol.* **12**, 5–24. (doi:10.1086/201401)
8 Rizzolatti, G. & Arbib, M. G. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194. (doi:10.1016/S0166-2236(98)01260-0)
9 Gibson, K. R. & Ingold, T. (eds) 1993 *Tools, language and cognition in human evolution.* Cambridge, UK: Cambridge University Press.
10 Arbib, M. A. 2005 From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**, 105–167.
11 Arbib, M. A. 2006 A sentence is to speech as what is to action? *Cortex* **42**, 507–514. (doi:10.1016/S0010-9452(08)70388-5)
12 Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)
13 Rizzolatti, G., Fadiga, L., Gallesc, V. & Fogassi, L. 1996 Premotor cortex and the recognition of motor actions. *Brain Res.* **3**, 131–141.
14 Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G. & Fogassi, L. 2008 Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *Eur. J. Neurosci.* **28**, 1569–1588. (doi:10.1111/j.1460-9568.2008.06395.x)
15 Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)
16 Buccino, G. *et al.* 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404.
17 Rizzolatti, G., Fogassi, L. & Gallese, V. 2009 The mirror neuron system: a motor-based mechanism for action and intention understanding. In *The Cognitive Neuroscience IV* (ed. M. Gazzaniga), pp. 625–640. Cambridge, MA: MIT Press.
18 Fogassi, L. & Ferrari, P. F. 2011 Mirror systems. *WIREs Cogn. Sci.* **2**, 22–38. (doi:10.1002/wcs.89)
19 Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. 1995 Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* **73**, 2608–2611.
20 Iacoboni, M. 2009 Imitation, empathy, and mirror neurons. *Annu. Rev. Psychol.* **60**, 653–670. (doi:10.1146/annurev.psych.60.110707.163604)
21 Pulvermüller, F. & Fadiga, L. 2010 Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* **11**, 351–360. (doi:10.1038/nrn2811)
22 Tobias, P. V. 1987 The brain of *Homo habilis*: a new level of organisation in cerebral evolution. *J. Hum. Evol.* **16**, 741–761. (doi:10.1016/0047-2484(87)90022-4)
23 Begun, D. & Walker, A. 1993 The endocast. In *The Nariokotome* Homo erectus *skeleton* (eds A. Walker & R. Leakey), pp. 326–358. Cambridge, MA: Harvard University Press.
24 Sherwood, C. C., Broadfield, D. C., Holloway, R. L., Gannon, P. J. & Hof, P. R. 2003 Variability of Broca's area homologue in African great apes: implications for language evolution. *Anat. Rec.* **271A**, 276–285. (doi:10.1002/ar.a.10046)
25 Schenker, N. M, Hopkins, W. D., Spocter, M. A., Garrison, A R., Stimpson, C. D., Erwin, J. M., Hof, P. R. & Sherwood, C. C. 2010 Broca's area homologue in chimpanzees (*Pan troglodytes*): probabilistic mapping, asymmetry, and comparison to humans. *Cereb. Cortex.* **20**, 730–742. (doi:10.1093/cercor/bhp138)
26 Petrides, M. & Pandya, D. N. 2009 Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* **7**, e1000170. (doi:10.1371/journal.pbio.1000170)
27 Rizzolatti, G. & Craighero, L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)
28 Ferrari, P. F., Gallese, V., Rizzolatti, G. & Fogassi, L. 2003 Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* **17**, 1703–1714. (doi:10.1046/j.1460-9568.2003.02601.x)
29 Fadiga, L. & Craighero, L. 2006 Hand actions and speech representation in Broca's area. *Cortex* **42**, 486–490. (doi:10.1016/S0010-9452(08)70383-6)
30 Gentilucci, M. & Corballis, M. C. 2006 From manual gesture to speech: a gradual transition. *Neurosci. Biobehav. Rev.* **30**, 949–960. (doi:10.1016/j.neubiorev.2006.02.004)
31 Roy, A. C. & Arbib, M. A. 2005 The syntactic motor system. *Gesture* **1**, 7–37. (doi:10.1075/gest.5.1.03roy)
32 Fiebach, C. J. & Schubotz, R. I. 2006 Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex* **42**, 499–502. (doi:10.1016/S0010-9452(08)70386-1)
33 Johnson-Frey, S. H. 2003 What's so special about human tool use? *Neuron* **39**, 201–204. (doi:10.1016/S0896-6273(03)00424-0)
34 Guiard, Y. 1987 Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *J. Motor Behav.* **19**, 486–517.
35 Hopkins, W. D. *et al.* 2011 Hand preferences for coordinated bimanual actions in 777 great apes: implications for the evolution of handedness in hominins. *J. Hum. Evol.* **60**, 605–611. (doi:10.1016/j.jhevol.2010.12.008)
36 Iriki, A., Tanaka, M. & Iwamura, Y. 1996 Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* **14**, 2325–2330.
37 Iriki, A. & Taoka, M. 2012 Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language

from the control of reaching actions. *Phil. Trans. R. Soc. B* **367**, 10–23. (doi:10.1098/rstb.2011.0190)

38 Macellini, S., Maranesi, M., Bonini, L., Simone, L., Rozzi, S., Ferrari, P. F. & Fogassi, L. 2012 Individual and social learning processes involved in the acquisition and generalization of tool use in macaques. *Phil. Trans. R. Soc. B* **367**, 24–36. (doi:10.1098/rstb.2011.0125)

39 Calvin, W. H. 1983 A stone's throw and its launch window: timing precision and its implications for language and hominid brains. *J. Theor. Biol.* **104**, 121–135. (doi:10.1016/0022-5193(83)90405-8)

40 Hopkins, W. D., Russell, J. L. & Schaeffer, J. A. 2012 The neural and cognitive correlates of aimed throwing in chimpanzees: a magnetic resonance image and behavioral study on a unique form of social tool use. *Phil. Trans. R. Soc. B* **367**, 37–47. (doi:10.1098/rstb.2011.0195)

41 Cantalupo, C. & Hopkins, W. D. 2010 The cerebellum and its contribution to complex tasks in higher primates: a comparative perspective. *Cortex* **46**, 821–830. (doi:10.1016/j.cortex.2009.10.004)

42 Boesch, C. 1993 Transmission of tool-use in wild chimpanzees. In *Tools, language and cognition in human evolution* (eds K. R. Gibson & T. Ingold), pp. 171–183. Cambridge, UK: Cambridge University Press.

43 Frey, S. H. & Povinelli, D. J. 2012 Comparative investigations of manual action representations: evidence that chimpanzees represent the costs of potential future actions involving tools. *Phil. Trans. R. Soc. B* **367**, 48–58. (doi:10.1098/rstb.2011.0189).

44 Hoffecker, J. 2007 Representation and recursion in the archaeological record. *J. Archaeol. Method Theory* **14**, 359–387. (doi:10.1007/s10816-007-9041-5)

45 Isaac, G. L. 1976 Stages of cultural elaboration in the Pleistocene: possible archaeological indicators of the development of language capabilities. *Ann. N. Y. Acad. Sci.* **280**, 275–288. (doi:10.1111/j.1749-6632.1976.tb25494.x)

46 Wynn, T. 1991 Tools, grammar, and the archaeology of cognition. *Camb. Archaeol. J.* **1**, 191–206. (doi:10.1017/S0959774300000354)

47 Bril, B., Smaers, J., Steele, J., Rein, R., Nonaka, T., Dietrich, G., Biryukova, E., Hirata, S. & Roux, V. 2012 Functional mastery of percussive technology in nut-cracking and stone-flaking actions: experimental comparison and implications for the evolution of the human brain. *Phil. Trans. R. Soc. B* **367**, 59–74. (doi:10.1098/rstb.2011.0147)

48 Stout, D. & Chaminade, T. 2012 Stone tools, language and the brain in human evolution. *Phil. Trans. R. Soc. B* **367**, 75–87. (doi:10.1098/rstb.2011.0099)

49 Ambrose, S. H. 2010 Coevolution of composite-tool technology, constructive memory, and language: implications for the evolution of modern human behavior. *Curr. Anthropol.* **51**, S135–S147.

50 Barney, A., Martelli, S., Serrurier, A. & Steele, J. 2012 Articulatory capacity of Neanderthals, a very recent and human-like fossil hominin. *Phil. Trans. R. Soc. B* **367**, 88–102. (doi:10.1098/rstb.2011.0259)

51 Pastra, K. & Aloimonos, Y. 2012 The minimalist grammar of action. *Phil. Trans. R. Soc. B* **367**, 103–117. (doi:10.1098/rstb.2011.0123)

52 Glenberg, A. M. & Gallese, V. In press. Action-based language: a theory of language acquisition, comprehension, and production. *Cortex*. (doi:10.1016/j.cortex.2011.04.010)

53 Tettamanti, M. & Moro, A. In press. Can syntax appear in a mirror (system)? *Cortex*. (doi:10.1016/j.cortex.2011.05.020)

54 Liebal, K. & Call, J. 2012 The origins of non-human primates' manual gestures. *Phil. Trans. R. Soc. B* **367**, 118–128. (doi:10.1098/rstb.2011.0044)

55 Call, J. & Tomasello, M. 2007 (eds) *The gestural communication of apes and monkeys*. Mahwah, NJ: Lawrence Erlbaum Associates.

56 Cartmill, E. A., Beilock, S. & Goldin-Meadow, S. 2012 A word in the hand: action, gesture, and mental representation in humans and non-human primates. *Phil. Trans. R. Soc. B* **367**, 129–143. (doi:10.1098/rstb.2011.0162)

57 Roby-Brami, A., Hermsdörfer, J., Roy, A. C. & Jacobs, S. 2012 A neuropsychological perspective on the link between language and praxis in modern humans. *Phil. Trans. R. Soc. B* **367**, 144–160. (doi:10.1098/rstb.2011.0122)

Review

Triadic (ecological, neural, cognitive) niche construction: a scenario of human brain evolution extrapolating tool use and language from the control of reaching actions

Atsushi Iriki* and Miki Taoka

*Laboratory for Symbolic Cognitive Development, RIKEN Brain Science Institute,
2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan*

Hominin evolution has involved a continuous process of addition of new kinds of cognitive capacity, including those relating to manufacture and use of tools and to the establishment of linguistic faculties. The dramatic expansion of the brain that accompanied additions of new functional areas would have supported such continuous evolution. Extended brain functions would have driven rapid and drastic changes in the hominin ecological niche, which in turn demanded further brain resources to adapt to it. In this way, humans have constructed a novel niche in each of the ecological, cognitive and neural domains, whose interactions accelerated their individual evolution through a process of *triadic niche construction*. Human higher cognitive activity can therefore be viewed holistically as one component in a terrestrial ecosystem. The brain's functional characteristics seem to play a key role in this triadic interaction. We advance a speculative argument about the origins of its neurobiological mechanisms, as an extension (with wider scope) of the evolutionary principles of adaptive function in the animal nervous system. The brain mechanisms that subserve tool use may bridge the gap between gesture and language—the site of such integration seems to be the parietal and extending opercular cortices.

Keywords: primates; parietal cortex; spatial integration; coordinate transformation; non-spatial cognition

1. INTRODUCTION

In the course of human evolution and human history, our ancestors have created new habitats from modified hunter–gatherer environments to agricultural landscapes with villages, and then to modern civilized technological cities. The evolution of various new cognitive capacities, including those underwriting the manufacture and use of tools and the production and comprehension of languages, has enabled these ecological transformations. Such new cognitive capacities in turn are an outcome of the dramatic expansion of the human brain and of new functional brain areas. Humans have constructed a new 'niche' in each of these ecological, cognitive and neural domains. 'Niche-construction' denotes an evolutionary process whereby the activities of organisms modify their habitat, to which in turn the organisms evolve to adapt, thus creating their own 'ecological niche' in the environment [1–3]. This concept will be extended in this paper to include the 'cognitive niche' as a newly acquired class of cognitive capacity [4], and the 'neural niche' as a portion of neural tissue added through expansion of the brain [5,6].

The above three classes of niche have coevolutionary interdependencies. Such interactions might have accelerated hominin evolution, which seems remarkably rapid if it was simply the product of natural selection driven by exogenous environmental change. It is possible that ecological changes to habitats have occurred not as a cause of hominin cognitive evolution, but rather as a result of it, with consequent selection pressures acting on the neural basis of behavioural adaptations to the modified environment. New brain functions would constitute the basis for further innovation in cognitive functioning and thus further modifications to the ecological niche, providing a feedback loop for 'triadic niche constructions'. In this paper, a potential evolutionary scenario that led humans to invent successively more complex forms of tools, and eventually to acquire the language faculty, will be proposed based on this dynamic interaction. The brain's functional characteristics play a key role in the above triadic interactions. The relevant neurobiological mechanisms are explored in this paper in order that the proposed evolutionary scenario should not be seen as teleological. Finally, we will try to locate these mechanisms as an extension of the

* Author for correspondence (iriki@brain.riken.jp).

evolution of the animal nervous system in general: human higher cognitive activity can then be viewed as continuous with that of other animal species comprising the wider terrestrial ecosystem.

2. CONSTRUCTION OF THE ECOLOGICAL NICHE
(a) *Hominin ecology structured through incorporating different classes of tools*
Humans are peculiar, compared with non-human species, in the extent to which they try to 'improve' their habitual environment. To make such improvements, particularly in our modern urban environment, we make and use various kinds of tools and technologies, and often the tools themselves are incorporated into the fundamental structure of the environment to create a distinctive human ecological niche. For example, cars running on paved roads or air-conditioned skyscrapers are essentially artefacts in which to travel and to reside, yet also comprise an environment to which city inhabitants adapt both physically and perceptually. In earlier times, hominins may also have typically constructed their niche by gradually and consecutively incorporating artefacts into the habitual environment. Thus, one can ask whether the tools that comprise our environment can be classified and structured in hierarchical order and if so, how different modes of brain function subserve their use.

The classical definition of the tool, as used in most existing tool-use studies, is restricted to external objects held by the hand and interacting with the external environments [7]. Namely, the tools so defined and studied are the ones that extend and externalize our hand, or more generally the motor organs or effectors. Indeed, the first series of tools that early hominins are known to have used [8,9], and those used by non-human animals [7] are these 'motor tools' (figure 1c, bottom row). We modern humans also use tools to extend or externalize our existing sensory organs, or to support the detection of information that is outside our natural sensory range (figure 1c, middle row). The optical telescope, endoscope or stethoscope would be examples of the former and the radio telescope or Geiger counter would be examples of the latter. Non-human animals rarely use this class of 'sensory tools' [10]. In monkeys, our own previous studies have demonstrated that they can be trained to use a sort of endoscope only after having acquired an ability to use a motor tool (a rake) [11]—as if attaching an additional visual cue to the tip of an extended body schema that was acquired through initial training to use a rake (see §2b for details). These results suggest that the class of sensory tools comprise a higher layer, superimposed onto previously acquired motor tools as the fundamental layer. Indeed, the history of our own technology suggests that sensory tools appeared much later, after motor tools were incorporated into human cultures.

What then would be the tool class of the third layer? If we looked at ourselves through our own externalized eyes (the second layer of the tools), we would observe ourselves as external objects by shifting from the first-person to the third-person perspective, in other words

by 'self-objectification' [12,13]. This leads to the perception of our own intrinsically intransitive movement as transitive, i.e. to the acquisition of a sense of the self (as the subject), and leading to the movement of ourselves or our body parts perceived as objects. We may hypothesize that once the 'self' has been bifurcated into a subjective self and an objective self, the mind and/or intentionality emerges as a function that bridges those fragmented 'selves' and reunites them; this hypothesis has been proposed in detail elsewhere [6,13]. As a result of this self-objectification and emergence of the 'mind', a recognition of the 'core self' that continues across time from the past through the present towards the future may subsequently arise. Once the future self is recognized as having a core that is identical to that of the present self, one might wish to save the present information for future use. This can be accomplished by taking notes or drawing pictures, which requires an external device for memorizing facts; thus, an 'externalization of the brain' is produced as the tool class of the third layer (figure 1c, top row) [6,14].

How could the three different classes of tools outlined above be incorporated into humans' habitual environments? The scenario outlined proposes that successive layers of tools (motor extensions, sensory extensions and symbolic externalized memory) can be incorporated into the environment by building on the pre-existing acquisition and incorporation of tool classes of the immediately lower level. Thus, a positive feedback would have emerged between new brain functions and resulting modifications of the habitual environment. In the course of such positive feedback processes, a brain function emerged for the mind and for future-directed 'intentionality', after which the feedback became guided by human intentions ('intentional niche construction') [6].

(b) *Parietal plasticity when incorporating tools into the body schema*
In using these tools, what kind of neural mechanisms and what modes of operation are employed, and what kinds of neural changes, if any, are induced upon acquisition of the ability to use tools? Our previous studies, as illustrated below, demonstrate one such example. Although Japanese macaques normally do not use tools in their natural habitat, two weeks of extensive training will enable these animals to use a hand-held rake to retrieve a distant food object located out of reach [15]. This training must imply the ability to reorganize the image of the body to one in which the rake is incorporated as an extension of the forearm. The body image is thought to form by integrating somatosensory and visual information relating to the body [16]. Thus, its modification after tool use could be physically observed as changes in the receptive field properties of the neurons that code such images [17]—when the tool was incorporated, the receptive field that codes the image of the hand was elongated to include the rake (figure 1a(iii)). This modification seemed to match the monkeys' internal states, whether or not the rake was incorporated into the image of the forearm. Here, an equivalence is established between body
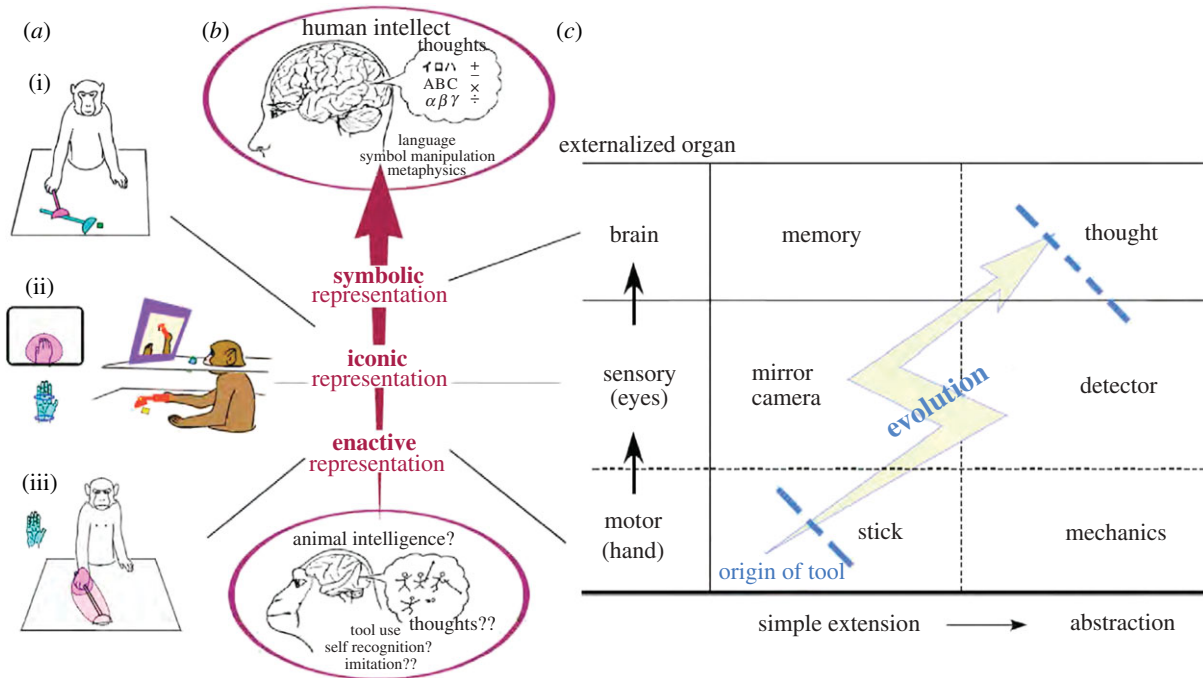
Figure 1. Various modes of cortical body-image codings (*a*) and hierarchical structure of various classes of tools (*c*) corresponding to the putative hierarchy of internal representations (*b*). (*a*(i)) Combinatory usage of short and long rakes. (*a*(ii)) When monkeys used a monitor, a visual receptive field of representative intraparietal bimodal neurons was formed around the hand in the monitor, encompassing its somatosensory receptive field. (*a*(iii)) When monkeys used a rake to retrieve distant food, the visual receptive field encompassed the somatosensory receptive field of a representative intraparietal bimodal neuron extended along the rake. Reproduced with permission from Iriki & Sakura [6].

parts (hands) and tools, i.e. hands are extended towards tools (externalization of the innate body) or tools are assimilated into the body schema (internalization of external objects). Hence, these representations of the body image comprise an internal model of the bodily structures used to control various movements, as a concrete neural correlate of the 'enactive representation' [18], a class of representation that first emerges at a very early stage of postnatal development in human infants (figure 1*b*).

As the training proceeds further, we might postulate that the monkey's mode of representation may advance to 'iconic (visual)' and even close to 'symbolic' [18]—during human development, these appear later during childhood or after maturity. This expectation implies that motor-tool-use-trained monkeys could be further trained to use a video monitor to retrieve food that is out of their direct line of sight, and that the receptive field of the parietal neurons that code the hand and the tool incorporated into it will be activated by visual feedback when images of the hand and tool are seen on the video monitor (figure 1*a*(ii)) [19]. Thus, the body image is visually projected onto the distant monitor screen. In fact, we have found that monkeys that acquire the ability to use a rake and a video monitor to retrieve food objects in this way can immediately combine multiple tools purposefully to accomplish the goal [20], as if they are able to logically structure body parts using (proto-)symbolic representations (figure 1*a*(i)). Here, we can recognize that the hierarchical structures of motor tools/sensory tools/brain tools [18] resemble the hierarchy of representations from enactive (motor) to iconic (sensory) to symbolic (brain) structures of development, as

described earlier (see figure 1*a–c* for these comparisons) [6]. Thus, reorganization of the modes of visuomotor integration in the parietal cortex must be crucial for the acquisition of these successively more advanced modes of representation.

3. CONSTRUCTION OF THE NEURAL NICHE
(a) *Brain expansion by tool-use training*
Is the neural plasticity depicted above limited within the range of the individual's learning capacity, or could it cumulatively evolve over generations? In other words, is it purely subserved by 'cultural inheritance', or alternatively, could it be a part of an epigenetic evolutionary mechanism in which the information embedded in the environment contributes to modification of phenotypic expression in succeeding generations? Although the latter has a flavour of Lamarckism—inheritance of acquired phenotypic traits—there may be a biological mechanism that could channel the evolution of adaptations to an environment in which cultural information is embedded. Macroscopic expansion (up to 23% of MRI grey matter signal) of cortical grey matter, including the intraparietal region, was detected in monkeys undergoing two weeks of tool-use training (figure 2*a*) in our recent Voxel Based Morphometry analysis (a kind of digital neuroanatomy using the magnetic resonance imaging technique) [21]. During the same period, microscopic changes (axonogenesis and synaptogenesis, as detected by tracer-injection histological examinations) [22] together with elevated expression of immediate-early genes [23] and of neurotrophic factors [24,25] were also shown to have been induced in these cortical
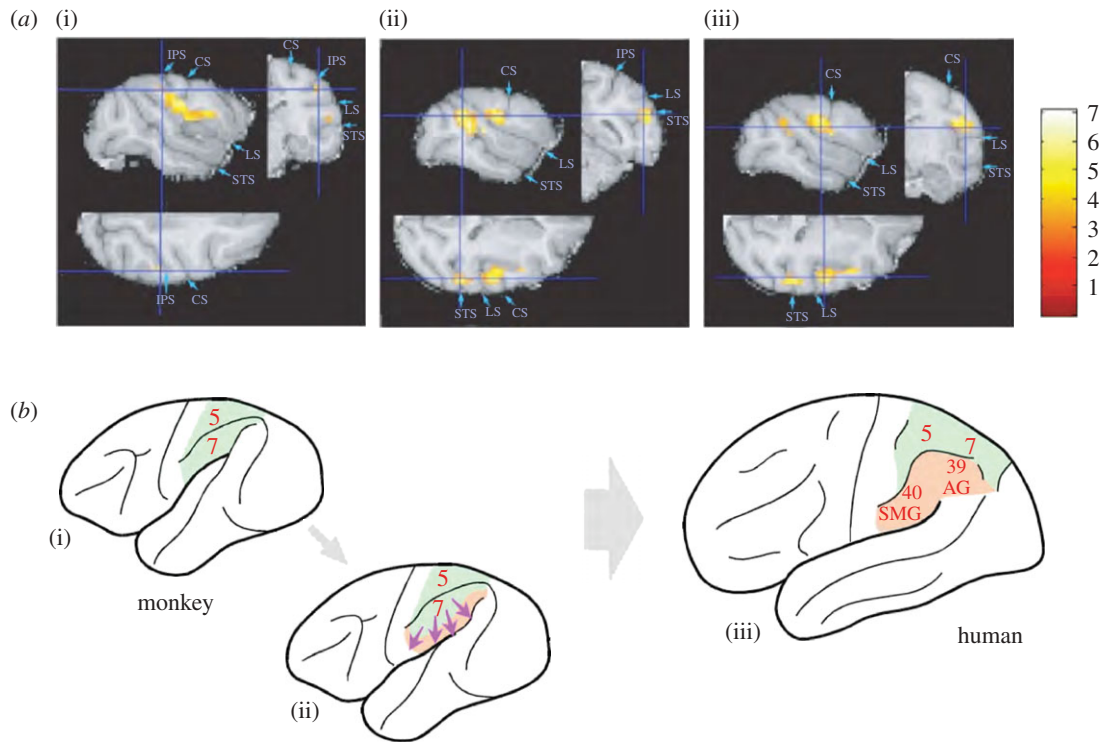
Figure 2. Grey matter increase with improvement in rake task performance. (*a*) Areas where grey matter increased with increasing performance score on the rake task. Sagittal, coronal and horizontal planes with increases in grey matter, including the right intraparietal sulcus (IPS, (i)), the superior temporal sulcus (STS, (ii)) and the secondary somatosensory area (SII, (iii)), are shown. (*b*) Schematic illustrating how tool-use-induced expansion of the parietal cortex of monkeys (i,ii) may contribute to the establishment of a precursor for the formation of human inferior parietal areas (iii), thus creating a novel neural niche that subserves further higher cognitive functions. CS, central sulcus; LS, lateral sulcus. The colour scale indicates the *t* score. (*a*) Reproduced with permission from Quallo *et al.* [21].

areas. The grey matter expansion extended to include adjacent areas, such as the secondary somatosensory area (figure 2*a*(iii)) and the surrounding opercular cortex. Although the evidence obtained to date remains fragmentary and more detailed biological examinations are in progress, these initial findings indicate that the brain is much more adaptive than was previously believed: exposure to a novel cultural environment induces the brain to exhibit not only functional plasticity, but also extensive and persistent morphological change.

This implies that once a novel cognitive demand, such as incorporation of motor tools into the body schema, has become embedded in the environment, modifications of brain structure would be induced automatically through the normal developmental processes in succeeding generations. The occurrence of such a plastic response during the lifespan as a result of behavioural modifications that lie within the existing adaptive capacity of individuals, and its subsequent consolidation (under selection acting on changing gene frequencies) as a default state that is stable over generations [26–28] is termed the 'Baldwin effect' [29,30], and comprises one potential component of the evolutionary process.

(b) Redundant and polysemic systems as pre-adaptations for a novel neural niche

What could be the biological principles that allow brain expansion as an evolutionary mechanism of the kind just outlined? Biological systems are never ultimately efficient—systems require some redundancy for stability, to avoid over-specialization that might threaten the capacity to survive new challenges. Some redundancy in brain structure would allow representational bistability, for both the originally adapted functions and functional response to the new challenge. Increased redundancy to stabilize this functional capacity for flexible adaptive responses, perhaps via rapid brain expansion, would also then allow rapid construction of new and specialized neural resources. As has been discussed earlier, monkey intraparietal neurons that normally code body image can be trained to code a tool in a way that is equivalent to that for the hand holding it [17]. Thus, these neurons are bistable or polysemous for the hand or the tool. This functional plasticity may be an inherent property at the margins of a neural coding system prepared for gradual elongation of the arm during body growth, and which can then also adapt to a 'sudden elongation' by using the tool. This accidentally established equivalence between body parts (hands) and tools in turn leads to additional polysemic and bistable *interpretations*, i.e. hands may be extended into the tool representation (externalization of the innate body) or tools may be assimilated into the body schema (internalization of external objects).

Thus, redundancy in the brain, initially adapted to stabilize this system against unexpected environmental noise (or developmental changes, following the above speculation about body growth) has occasionally

allowed the system to be polysemous. This newly acquired bistable state enables the reuse of cortical systems for different functions in the future, as in the case of tool use, perhaps in combination with other parts of the brain [5,31]. This bistability, or 'polysemy', could enable the use of metaphors in conceptual structure—so as to comprise a novel cognitive niche, as will be described in the next section. However, how human higher cognitive functions appear to have 'evolved' much more quickly than might be expected from ordinary biological evolutionary processes of adaptation to changing external environmental contingencies remains an open question. Humans can induce such changes intentionally, to construct a better-fitting ecological niche [6]. The capacity of human intention, or of the human mind, to plan for the future emerged through the process described in the previous section in relation to the hierarchy of classes of tool use (motor, sensory and brain). Subsequently, the neural systems which process the information that is necessary to inhabit the tool-modified environment (the neural niche of the brain; figure 2b(ii)) could be reinforced further by extragenetic or epigenetic triggering factors embedded in such an environment.

In addition, because hominin species have attained an unusually long post-reproductive lifespan, particularly females [32], accumulation of knowledge continues over the whole lifespan of an individual, tending to peak in middle-to-old age. Thus, some extragenetic mechanisms are indispensable for inheritance of these later acquired 'cognitive niches' over generations to occur. Species with a short lifespan and mass reproduction adapt to environmental changes through variations in their numerous offspring, as they expect that at least a few will survive, whereas species with a long lifespan—such as primates, and most typically humans—and low birth rate do so through an individual capacity to adapt [33]. This process would be enhanced by expansion of an organ that controls adaptive behaviours, namely the brain, which are stabilized as the typical phenotype of the species through epigenetic mechanisms. The evolutionary process driving such expansion may not simply be natural selection acting on random mutations, but rather something like the Baldwin effect [29,30] as depicted earlier.

4. CONSTRUCTION OF THE COGNITIVE NICHE
(a) *Parietal 'neural niches' for processing spatial and non-spatial cognition*
Debate exists on the comparative anatomy of the primate parietal cortex (figure 2b). One view claims that the inferior parietal area is evolutionarily new and uniquely expanded in humans, and that therefore the monkey posterior parietal cortex corresponds to the human superior parietal lobule. This is the position illustrated in the scheme presented in figure 2b(iii), independent of figure 2b(i) [34,35]. Evidence supporting this view would include the fact that, for example, the superior parietal lobule tends to process spatial information in a conventional way, whereas the inferior parietal lobule is often credited

with non-spatial cognition. Many recent human imaging studies demonstrate that this cortical area additionally supports various forms of high-order non-spatial cognition that are not necessarily directly related to physical space itself. Indeed, a comparison of the brain areas responsible for tool-use behaviours in monkeys and humans [36] detected a patch of the parietal cortex specific to humans that might be responsible for perception and manipulation of abstract causal relationships required for human tool-use behaviours. This could be evidence of a function that was derived from a polysemic mechanism as described above, in which an additional resource of brain tissue (a neural niche) functions to enable an additional cognitive process (a cognitive niche).

However, there is another view in which the monkey parietal cortex includes functional homologues of both regions [37–40], which allows viewing figure 2b(i–iii) as a continuum. Various kinds of non-spatial cognition can be grouped and ordered based on the levels of abstraction of the 'objects' and the conceptually defined spaces which are represented. The assumed coordinate systems for such 'spaces', with citations to research literature analysing the brain mechanisms which encode such spatial coordinates (for monkeys and humans), are summarized in figure 3 [108]. The pseudo-spatial nature of the high-order cognition supported by the posterior parietal cortex may be derived from the essential characteristics of the objects represented; alternatively, it may be derived from the nature of the pre-existing information-processing mechanisms of this area, namely as a hub for multi-sensory integration and for representing physical environmental space [109]. A meta-analysis of the references listed in figure 3 (refer to its legend for a detailed classification of various cognitive niches handled by this brain area) shows that the posterior parietal areas responsible for these novel forms of cognition are not necessarily clearly segregated, either in monkeys or in humans, and suggests a trend of gradual expansion towards the lateral sulcus as the level of abstraction increases (figure 3c) [108]. Thus, it seems that the parietal area gradually incorporated high-order cognition as it expanded during hominid evolution, while preserving its original principles of operation. This could be an example of exhibiting a novel cognitive niche by reusing the functions that have derived from a polysemic mechanism described previously. If this was the case, the gradual emergence and differentiation of functions in the transitional state might not be detected until the intensity of activation and the quantity of tissue recruited to serve this new neural niche exceeded some threshold of detection. In this sense, the above two views of parietal evolution are not necessarily mutually exclusive.

What, then, is the explanation for this series of additions of novel cognitive niches? Mechanisms for 'selecting' and 'switching' between objects among different represented spaces (figure 3a) could be hypothesized to contribute to this process [108]. That is, initially in classical (physical) space, spatial attention towards concrete objects was typically expressed as the direction of the eye axis to represent
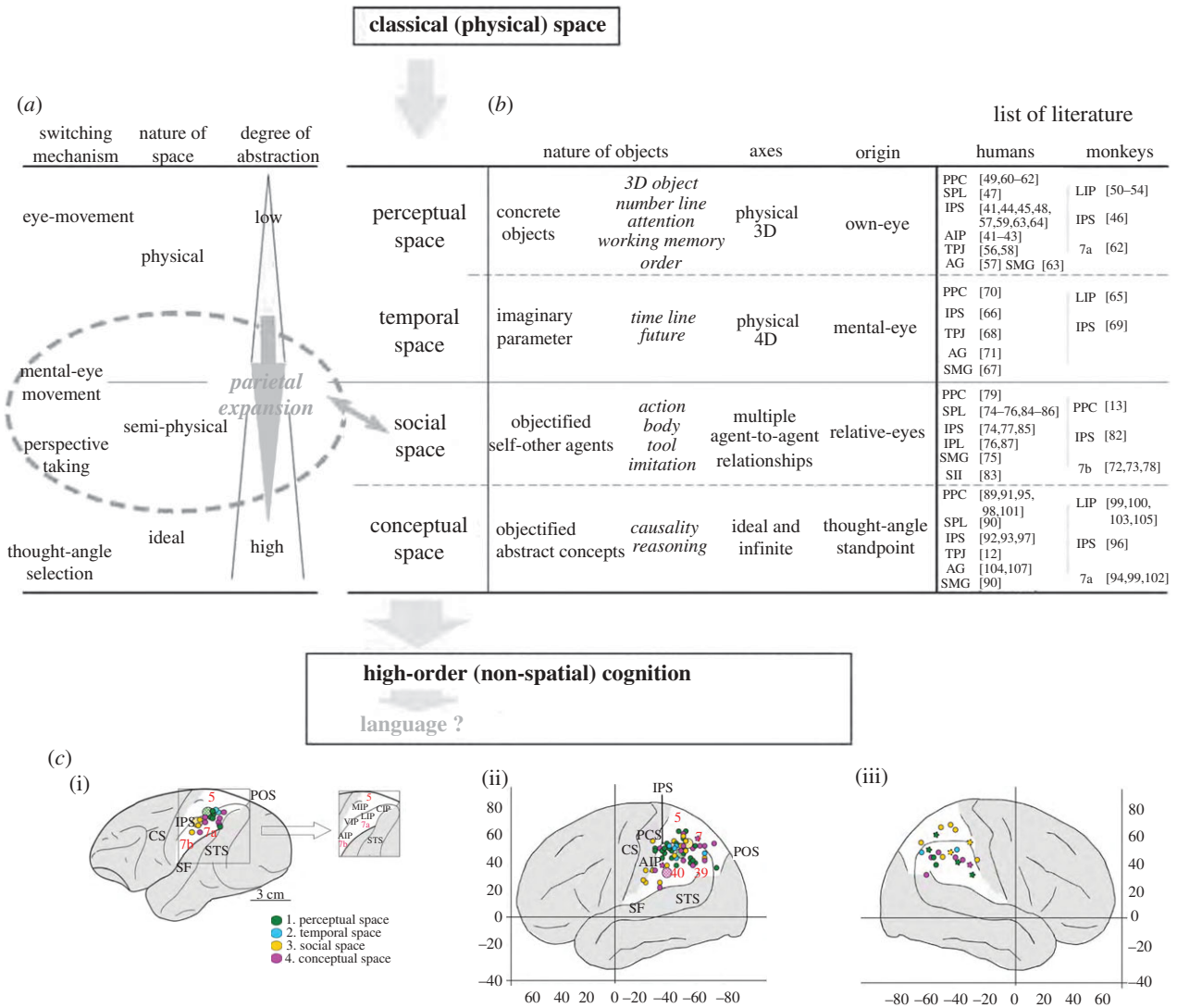
Figure 3. (*Caption opposite.*)

'perceptual space' ([41−64]; see legend of figure 3 for details of respective references, and also for other classes of 'spaces' described below). Secondarily, when such attention had to be sustained or the attending content had to be memorized, invisible 'time' was 'visualized' in the mind's eye, becoming a new virtual dimension in the existing suite of spatial-coordinate systems, namely 'temporal space' [65−71]. And then, once one was able to visualize an invisible virtual entity, a similar objectification process could have been extended further, enabling intentional perspective switching. Acquiring representations of 'social space' [13,72−87] might have accelerated this process. Via self-objectification processes [13] mentioned in the earlier section, and the development of 'virtual eyes' [11,12], flexible and mutually integrated representations of the bodily self, of the analogous selves of others, and of tools used as equivalents of body parts (and vice versa) may have served as a bridge between concrete physical and abstract conceptual spaces. Finally, as the posterior parietal cortex expanded in both physical volume and in range of function [6,21], a positive feedback process could have been established to achieve further human-specific forms of non-spatial conceptual cognition, or 'conceptual space' [12,80,81,88−107]. In this way, crucial components of human intelligence would derive their character from the precursory spatial cognition process of the parietal cortex. Language is full of spatial metaphors for abstract thoughts.

(b) Opercular cortex as a cradle for language by re-using spatial processing principles

During evolution, whenever organisms are faced with a novel and unforeseen environment, they have no other means to overcome immediate problems than to reuse any materials at hand [5,31]. Thus, cognitive capacities are extended by diverting pre-existing functions. In hominin evolution, according to the scenario outlined in the previous section, the expanded inferior parietal area and surrounding opercular region have taken on distinctive functional characteristics. Basic continuity from monkeys to humans as described above seems to be present in this general area, which includes Broca's area (anterior operculum), Wernicke's area (posterior operculum) and the middle operculum corresponding to the supramarginal and angular gyri, and which appear to be an extension of the inferior parietal lobule of the monkey brain; the continuity

Figure 3. (*Opposite.*) Non-spatial representations of the parietal cortex as a function of abstractness. (*a*) Hypothesized mechanism to increase the level of abstract representation in the parietal cortex. (*b*) Structure of various levels of abstract spaces. Existing references classified into each category are listed in the right-most columns (for humans and monkeys). Concrete functions analysed in each reference are as summarized below: (i) Perceptual space: the posterior parietal cortex (PPC) is implicated in three-dimensional object recognition, processing of number and quantity, attention, and memory. The intraparietal sulcus (IPS) and anterior intraparietal area are activated when concrete three-dimensional objects are recognized, explored, imagined and constructed [41,42]. Discerning continuous quantity recruits the anterior IPS [43], while discrete number processing recruits the bilateral IPS [44–46], forming a mental number line [47] and mental arithmetic [48]. Sustained attention to spatial locations recruits the area between the right IPS and the inferior parietal lobule (IPL) [49]. The lateral intraparietal area responds to various features of the stimulus shape [50–55]. Activity of angular gyrus and temporoparietal junction (TPJ) is related to stimulus saliency detection and control [56,57]. The left and right posterior TPJ are active while processing global and local information, respectively [58]. The bilateral IPS supports episodic and semantic memory [59]. The parietal cortex may play a critical role in working memory [60–63], although this region can also be active during tasks requiring no working memory [64]. Thus, the apparent role of this cortical area in working memory may reflect a broader function, such as temporally transferring information from present to future. (ii) Temporal space: the posterior parietal cortex is implicated in temporal processing [65,66], time estimation and future planning [67]. The TPJ plays a crucial role in temporal order judgement tasks [68] as well as in magnitude judgements about numbers [48]. The parietal cortex supports time-interval estimation and is active when estimating future object position [69,70]. Here, events are processed in order and placed in a timeline [71]. (iii) Social space: the parietal cortex processes some aspects of social space, including action understanding, tool use and self–other relationships, thereby enabling imitation. This is one of the brain areas showing mirror properties [72–74]. Gesture imitation activates the IPL and supramarginal gyrus (SMG) [75]. Imitation activates the left inferior parietal cortex more than observation does, regardless of first- or third-person perspective [76]. The human anterior IPS distinguishes between observed and executed movements [77], suggesting that parietal neurons have hierarchical properties for recognizing similarity or difference of kinematics, goals and function of motion [78]. The parietal cortex supports tool usage, an extension of the bodily self [79–82]. This faculty may require an implicit equivalence to be drawn between innate body parts and external objects. A proposed brain network subserving such body-part objectification includes the posterior parietal cortex [13]. This area supports the recognition and drawing of distinctions between agents [83,84]. Evaluations of physical and social distance to others involve the superior and inferior parietal cortices, respectively [85], in accordance with egocentric distance estimation by the parietal cortex [86]. Cooperative and competitive tasks recruit a common set of brain regions with competition activating the right IPL more strongly, reflecting the apparent contrast between self and others [87]. The facility with both concrete spatial cognition and self–other discrimination may have preadapted it to handle perspective transformation in social situations. This in turn could have laid the basis for further functional expansion into realms such as social categorization and, later in evolution, conceptual spaces [78]. (iv) Conceptual space: finally, this area appears to play a role in causal reasoning, mental object manipulation, attention switching, set shifting and dimensional abstraction. In addition to its role in tool usage, different areas support broad concepts such as the causal ramifications of using tools [81,88,89] and the logical structures of the gestures involved [80]. The superior parietal lobule (SPL) and SMG are differentially activated during rotations of visual and motor imagery, respectively [90]. Spatial scanning through mental imagery activated the precuneus, SPL, IPL and IPS [91]. The PPC is responsible for mentally manipulating sequentially learned materials [92], whereas IPS and SPL together with the mid-dorsolateral prefrontal cortex is responsible for monitoring the learned items [93]. PPC activation is also related to cognitive set shifting [94–99]. During a visual search with attention shifts, PPC neurons depend on a targeted dimension irrespective of the stimuli's spatial features [100]. PPC activation was diminished in elderly people in the solution-search phase of a set-shifting task [101]. Perspective taking [12] is one of the prominent qualitative distinctions between the attention-shifting abilities of humans and other primates [78]. Posterior parietal activity also correlates with abstract information at various levels and dimensions [102–105] and with modulation of the strength of the visual stimulus and motor planning [106]. The left parietal cortex is involved in the cortical rearranging of the relationship between semantic items in space [107]. (*c*) (i) Monkey posterior parietal cortex; (ii) human left posterior parietal cortex; (iii) human right posterior parietal cortex. Results of meta-analysis (based on the references listed in *b*) of cortical areas responsible for non-spatial cognition are superimposed. For those references which detected multiple activations, the strongest representative regions relevant to the functions discussed were plotted. Green, blue, yellow, pink symbols, respectively, indicate perceptual, temporal, social and conceptual representations. Data points in monkeys are projected onto the left hemisphere (because no particular laterality has been claimed), whereas those in human subjects are illustrated independently for each hemisphere (being biased towards the left hemisphere). Large and dashed symbols indicate areas estimated from reviews and original papers, respectively, in which coordinates were not clearly specified. Arabic numerals indicate Brodmann areas (40, supramarginal gyrus: SMG; 39, angular gyrus: AG). AIP, anterior intraparietal area; CIP, caudal intraparietal area; CS, central sulcus; IPS, intraparietal sulcus; LIP, lateral intraparietal area; MIP, medial intraparietal area; PCS, post central sulcus; POS, parieto-occipital sulcus; PPC, posterior parietal cortex; SF, sylvian fissure; STS, superior temporal sulcus; VIP, ventral intraparietal area; TPJ, temporoparietal junction (adapted from figures and text originally published in *F1000 Biology Reports*, 2009 Yamazaki *et al.* [108]).

underwrites opercular language representations. Fundamentally, the argument from continuity implies that such representational capacity should be a simple extension of a coding system for reaching and grasping. Initially, this extension would derive from the coding of spatial integration and of a reorganized representation of space, which could be extrapolated further using a principle identical to the non-spatial higher-order coding of more abstract objects [108,110]. In particular, this extrapolation could be subserved by the 'abstraction' of free and unconstrained polysemous handling of the space and of the body, which would comprise a fundamental component of language representations, and perhaps also by common neural mechanisms that share a mode of information integration and processing.

Human-specific illogical cognitive biases for symmetrical inference (the tendency to incorrectly infer 'if B then A' from a conditional relationship 'if A then B') and for inference by exclusion (the tendency spontaneously to assume that an unfamiliar label goes with an unfamiliar object) involve these same brain areas [111,112]. The mind, human language and human cultural transmission, all of which contribute to the semantic inheritance of the benefits acquired during the unusually elongated human post-reproductive lifespan, are aspects of cognitive functions that have evolved recently and result from such neural niche construction. Once a fundamental syntactic SVO (subject/verb/object) structure emerges [113] and is generalized, abstraction and concept formation and their manipulation become possible and constitute a basis for further intellectual advancements, such as polysemic interpretation of phenomena (which enables metaphorical inference). Such redundancy and polysemous representation would allow equivalence and symmetric inferences and would lead to the emergence of symbols. All of these functions seem to be carried out in the expanded inferior parietal and surrounding areas in humans. Hence, the human language faculty seems to draw on these fundamental neural mechanisms, which are found in these late-myelinating brain areas, which retain a large degree of flexibility until adulthood.

5. PARIETAL CORTEX AS THE CENTRE OF TRIADIC INTERACTIONS

(a) *A site for multiple sensory and motor integrations and coordinate transformations*

The posterior parietal cortex plays a central role in multi-sensory integration and recognition of environmental space. Such integration provides a basis for the production of movements of various body parts, including eyes, hand-arm, head and whole body through transformations between different coordinate systems. The principle of neural reuse [5,31], as depicted above, seems to apply here in enabling the evolution of higher cognitive functions and thus of human cultural niches. Once these fundamentals were established in the parietal cortex, prefrontal cortex could have developed further so as to use the information for further executive functions involving working memory and syntactic operations, which are

often argued to have been crucial for the evolution of human intelligence [114]. How, then, could this function have been extrapolated from the general evolution of the nervous system? In this section, we shall sketch a possible sequence of grades of gradually increasing complexity that might take us from reaching movements to tool use and language.

Throughout evolution from primitive protozoa to mammals, the mouth was the organ used both to grasp and to intake prey, after reaching it through locomotion along the body axis. Some animals, especially primates, finally developed the hand to reach and grasp. The target to reach (prey) is detected by sensory organs of various modalities. The nervous system links these sensory and motor apparatuses to produce appropriate actions. The site of such integration within the neuraxis expanded continuously, to finally form the parietal cortex in primates; and its further extrapolation enabled the use of tools, as depicted above. The emergence of bipedalism constituted another evolutionary path for such expansion, as it differentiated the body axis from the movement axis, thus demanding a dramatic increase in spatial information transformation. In turn, this drove the evolution of the parietal operculum in the neural niche. Figure 4 illustrates this scenario, of which fundamental behavioural and neural correlates are summarized below.

(i) *Head reaching*

Primitive animals, in which the locomotor apparatus (such as fins or limbs) has not yet evolved, 'crawl' with the whole body to prey (figure 4a). The 'mouth' is located at the front end of the body axis [115,116], where sensory organs cluster to efficiently acquire environmental information. Moving in the direction of the mouth (i.e. 'head reaching') is still common in extant taxa, including vertebrates. Fish swim in the water three-dimensionally along the body axis (figure 4b). Terrestrial amphibians (and essentially also reptiles and most mammalian species) were constrained to move two-dimensionally on the land surface, yet they still move in the direction of the main body axis and crawl to reach, having evolved limbs for locomotion and resistance against gravity (figure 4c). Head reaching requires information on the target from a self-centred perspective. Animals align the body axis (the direction to move) towards the object and then approach it by travelling with the whole body until arriving at the target. The neural machinery used need only be the rather stereotyped projection of the body onto the environmental space, which requires minimal resources of neural tissue, and of which even insects' tiny brains are capable [117,118].

(ii) *Neck reaching*

Avians further developed, from forelimbs, the wings to fly. After a flight to reach prey, two final precise reaching-and-grasping procedures emerged. Raptors grasp the prey object by a hind limb [119] and finally eat it using the mouth/beak—they use organs other than the mouth just to grasp, but not directly to eat. Species with long flexible necks, like herons or cranes, reach and grasp directly with the beak using neck
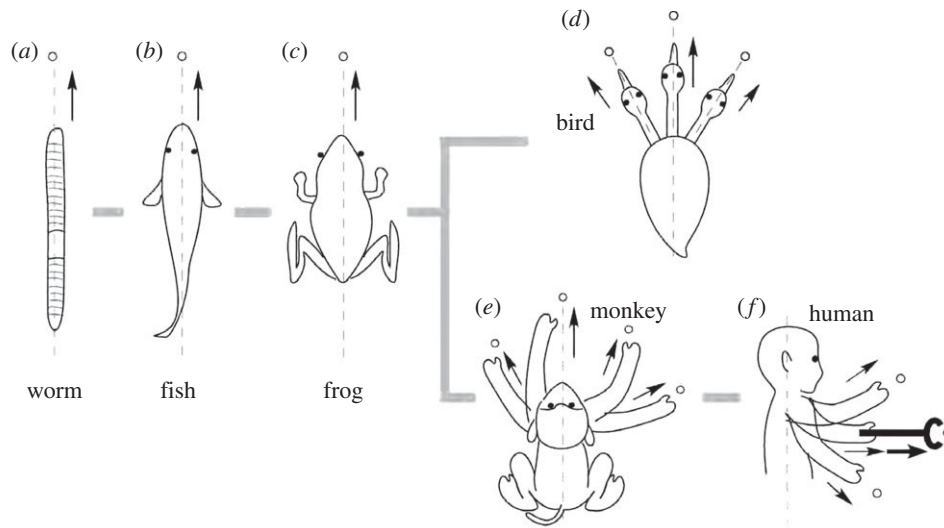
Figure 4. Patterns of reaching-and-grasping movements in different animal species. Fundamentally, animals reach and grasp by moving towards the targets (open circles) using their whole body (*a*–*c*). Higher vertebrates developed additional organs, elongated from the trunk to reach and grasp (birds (*d*, neck) and primates (*e*, forearm)). (*f*) In humans, the direction of the movements became perpendicular to the axis of the body because of bipedalism, and was further extended or transformed by the use of tools. Note the differences among the different species in the relative location of the eyes (and of the vestibule) relative to the organs used to finally reach and grasp, as well as their movements relative to the trunk and to the axis of the body (dashed lines).

movements [120,121]. Here, a discrepancy emerges between the axes of the body and the head (figure 4*d*). As long as the mouth and eyes remain relatively fixed, the neural processing to reach with the mouth/head remains the same. But, once the neck can move independently of the trunk, object locations need to be represented in multiple spatial-coordinate systems—not only for the original coordinates with the body/trunk at the origin, but also with the head moving relative to the body. Such transformations between coordinate systems would have required their brains to evolve further neural resources.

(iii) *Arm reaching*
In mammals, particularly primates, the forelimbs have evolved as apparatuses to reach and grasp, diverging from their original locomotor function. Such evolution occurred via (i) substantial elongation of the forelimbs; (ii) increased degrees of freedom of movement at the shoulder, elbow and wrist joints; and (iii) elongation of digits to grasp objects of various sizes, shapes and orientations [122–124]. These changes dramatically increased the diversity of kinds and orientations of reaching-and-grasping motions in the space around the body axis (figure 4*e*). However, as a trade-off, it requires complex information processing by the brain to harmonize the movement of different body parts by translating positional information between different coordinate systems—body-centred, eye–head-centred and hand-centred systems [125]. Such situations demand more neural resources and the evolution of highly developed spatial perception, resulting in the expansion of parietal cortex. This was a cradle for the further evolution of transformations and modifications between coordinate systems, even for other working spaces. These served as a preadaptation by increasing the degrees of freedom for spatial

information acquisition, thereby enabling further expansion of the brain areas that are responsible for those calculations.

(iv) *Rotation of moving axis*
These processes could be immediately extrapolated onto further evolutionary events. One of those would be the emergence of constant bipedalism. This consists in the maintenance of an upright head-lifted posture, with locomotion perpendicular to the axis of the body trunk for the first time in evolution (figure 4*f*). Additional constraints emerged, i.e. visual axes became fixed to the direction of locomotion (horizontal), thus also perpendicular to the axis of the body. As a result, various axes (body, hand, head and eyes) became dissociated and the directions of locomotion and of reaching/grasping became independent, depending on the ongoing behaviour. The brain mechanisms for processing such information remain incompletely understood, and open for future investigations.

(b) *Extension of axes from concrete to virtual spaces for locomotion, tool use and language*
In evolution, the parietal cortex expanded initially as an adaptation to demands from the environment, perhaps for control of different movements of the various body parts, while prefrontal cortex may have expanded later to control such information coded in the parietal cortex through prefronto–parietal interactions [114]. In this way, neural mechanisms became embedded in the brain which served as pre-adaptations for further neural evolution through neural reuse, ultimately enabling the language faculties and (via prefrontal expansion) advanced modes of executive control. The emergence of bipedalism, through its associated demand on multi-sensory integration and very complex sensorimotor coordinate transformations, also

pressurized expansion of the greater opercular regions of the cerebral cortex, and thereby facilitated subsequent reuse of such structures for higher cognitive functions including language.

Neural evolution in the first stage of this extension of axes put in place conditions for the emergence of tool use (derived from the usage of innate organs for a purpose not originally planned for). This polysemous pattern of organ usage, and the discovery of novel types of usage, would be key elements in inducing the re-evaluation of existing spatial structures in relation to the body axis as a second stage. The emergence of bipedalism triggered further extrapolation of such faculties and initiated the usage of the externalized body, i.e. extension of body parts into the tool representation. Such freedom from existing physical and bodily constraints in the understanding of the environmental space would allow a novel mode of spatial perception using novel tools (perspective transformation) that would be the basis for the next jump in the acquisition of abstract and transcendental thoughts—stage three. This development has served, in a final stage, as a cradle for the language faculty, principally by developing its neural basis for information processing, both for the use of polysemous and conceptual thoughts and for the articulation of oro-facial organs, to finally subserve language.

6. CONCLUSION

Expansion (or increase in capacity) of organs as an adaptive response to ecological pressures seems to be a general biological and evolutionary tendency to make the phenotypic system robust—the brain will not be an exception. Multi-sensory integration and coordinate transformation for the control of reaching movements in the inhabited space is an essential function of the nervous system, for which evolution finally endowed primates with a well-developed parietal cortex. The shift of body-space structure associated with the emergence of hominin bipedalism may have further pushed this trend forward to give this area, and the extended opercular cortex, further resources. Such neural enhancement (construction of the neural niche) happened to enable the processing of abstract information, detached from actual physical constraint, by applying and re-using existing principles for spatial information processing to realize novel mental functions (construction of the cognitive niche)—ultimately leading to language. Purposeful manipulation of the body image in space, required for tool use, would have accelerated interactive links between the neural and cognitive niches—tool use requires transformation of various bodily and spatial coordinates, as well as logical and sequential relations of action components.

Tools represent materialized cognitive brain functions. They have been created one after another and incorporated into hominin habitats as constituent elements (construction of the ecological niche). A human-modified environment puts pressure on succeeding generations to adapt to it, perhaps by acquiring further resources for the relevant organs. Epigenetically induced plasticity (including

developmental or learning mechanisms) would participate in such processes—and this is a subject for future biological investigations. In this way, extra genomic information could be transmitted between generations via mutual interactions among ecological, neural and cognitive domains of niches, which may have contributed to hominin evolutionary processes (that is, 'triadic niche construction'). This scenario would locate the human brain as part of an evolving holistic ecosystem.

REFERENCES

1 Kendal, J., Tehrani, J. J. & Odling-Smee, J. 2011 Human niche construction in interdisciplinary focus. *Phil. Trans. R. Soc. B* **366**, 785–792. (doi:10.1098/rstb.2010.0306)

2 Laland, K. N. & Brown, G. R. 2002 *Sense and nonsense: evolutionary perspectives on human behaviour*. Oxford, UK: Oxford University Press.

3 Odling-Smee, F. J., Laland, K. N. & Feldman, M. W. 2003 *Niche construction: the neglected process in evolution*. Princeton, NJ: Princeton University Press.

4 Pinker, S. 2010 Colloquium paper: the cognitive niche: coevolution of intelligence, sociality, and language. *Proc. Natl Acad. Sci. USA* **107**(Suppl. 2), 8993–8999. (doi:10.1073/pnas.0914630107)

5 Anderson, M. L. 2010 Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* **33**, 245–313. (doi:10.1017/S0140525X10000853)

6 Iriki, A. & Sakura, O. 2008 The neuroscience of primate intellectual evolution: natural selection and passive and intentional niche construction. *Phil. Trans. R. Soc. B* **363**, 2229–2241. (doi:10.1098/rstb.2008.2274)

7 Beck, B. B. 1980 *Animal tool behavior*. New York, NY: Garland.

8 de la Torre, I. 2011 The origins of stone tool technology in Africa: a historical perspective. *Phil. Trans. R. Soc. B* **366**, 1028–1037. (doi:10.1098/rstb.2010.0350)

9 Shennan, S. 2002 *Genes, memes and human history: Darwinian archeology and cultural evolution*. London, UK: Thames & Hudson.

10 Asano, T. 1994 Tool using behavior and language in primates. In *Behavior analysis of language and cognition* (eds S. C. Hayes, L. J. Hayes, M. Sato & K. Ono), pp. 145–148. Reno, NV: Context Press.

11 Yamazaki, Y., Namba, H. & Iriki, A. 2009 Acquisition of an externalized eye by Japanese monkeys. *Exp. Brain Res.* **194**, 131–142. (doi:10.1007/s00221-008-1677-1)

12 Corradi-Dell'acqua, C., Ueno, K., Ogawa, A., Cheng, K., Rumiati, R. I. & Iriki, A. 2008 Effects of shifting perspective of the self: an fMRI study. *Neuroimage* **40**, 1902–1911. (doi:10.1016/j.neuroimage.2007.12.062)

13 Iriki, A. 2006 The neural origins and implications of imitation, mirror neurons and tool use. *Curr. Opin. Neurobiol.* **16**, 660–667. (doi:10.1016/j.conb.2006.10.008)

14 Goldenberg, G. & Iriki, A. 2007 From sticks to coffee-maker: mastery of tools and technology by human and non-human primates. *Cortex* **43**, 285–288. (doi:10.1016/S0010-9452(08)70454-4)

15 Ishibashi, H., Hihara, S. & Iriki, A. 2000 Acquisition and development of monkey tool-use: behavioral and

kinematic analyses. *Can. J. Physiol. Pharmacol.* **78**, 958–966. (doi:10.1139/y00-063)

16 Head, H. & Holmes, G. 1911 Sensory disturbances from cerebral lesions. *Brain* **34**, 102–254. (doi:10.1093/brain/34.2-3.102)

17 Iriki, A., Tanaka, M. & Iwamura, Y. 1996 Attention-induced neuronal activity in the monkey somatosensory cortex revealed by pupillometrics. *Neurosci. Res. (NY)* **25**, 173–181.

18 Bruner, J. S., Olver, R. R. & Greenfield, P. M. 1966 *Studies in cognitive growth.* New York, NY: Wiley.

19 Iriki, A., Tanaka, M., Obayashi, S. & Iwamura, Y. 2001 Self-images in the video monitor coded by monkey intraparietal neurons. *Neurosci. Res. (NY)* **40**, 163–173.

20 Hihara, S., Obayashi, S., Tanaka, M. & Iriki, A. 2003 Rapid learning of sequential tool use by macaque monkeys. *Physiol. Behav.* **78**, 427–434. (doi:10.1016/S0031-9384(02)01006-5)

21 Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K., Lemon, R. N. & Iriki, A. 2009 Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proc. Natl Acad. Sci. USA* **106**, 18 379–18 384. (doi:10.1073/pnas.0909751106)

22 Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., Ojima, H., Obayashi, S., Fujii, N. & Iriki, A. 2006 Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. *Neuropsychologia* **44**, 2636–2646. (doi:10.1016/j.neuropsychologia.2005.11.020)

23 Iriki, A. 2005 A prototype of Homo faber: a silent precursor of human intelligence in the tool-using monkey brain. In *From monkey brain to human brain: a Fyssen Foundation symposium* (eds S. Dehaene, J.-R. Duhamel, M. D. Hauser & G. Rizzolatti), pp. 253–271. Cambridge, MA: MIT Press.

24 Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T. & Iriki, A. 2002 Tool-use learning selectively induces expression of brain-derived neurotrophic factor, its receptor trkB, and neurotrophin 3 in the intraparietal multisensory cortex of monkeys. *Brain Res. Cogn. Brain Res.* **14**, 3–9. (doi:10.1016/S0926-6410(02)00056-3)

25 Ishibashi, H., Hihara, S., Takahashi, M., Heike, T., Yokota, T. & Iriki, A. 2002 Tool-use learning induces BDNF expression in a selective portion of monkey anterior parietal cortex. *Brain Res. Mol. Brain Res.* **102**, 110–112. (doi:10.1016/S0169-328X(02)00201-2)

26 Jablonka, E. & Lamb, M. J. 2007 Precis of evolution in four dimensions. *Behav. Brain Sci.* **30**, 353–389.

27 Pigliucci, M. 2010 Genotype-phenotype mapping and the end of the 'genes as blueprint' metaphor. *Phil. Trans. R. Soc. B* **365**, 557–566. (doi:10.1098/rstb.2009.0241)

28 Tennie, C., Call, J. & Tomasello, M. 2009 Ratcheting up the ratchet: on the evolution of cumulative culture. *Phil. Trans. R. Soc. B* **364**, 2405–2415. (doi:10.1098/rstb.2009.0052)

29 Baldwin, J. M. 1896 A new factor in evolution. *Am. Nat.* **30**, 441–451. (doi:10.1086/276408)

30 Bateson, P. 2004 The active role of behaviour in evolution. *Biol. Phil.* **19**, 283–298. (doi:10.1023/B:BIPH.0000024468.12161.83)

31 Dehaene, S. & Cohen, L. 2007 Cultural recycling of cortical maps. *Neuron* **56**, 384–398. (doi:10.1016/j.neuron.2007.10.004)

32 Harman, S. M. & Talbert, G. B. 1985 Reproductive aging. In *Handbook of the biology of aging* (eds C. E. Finch & L. Hayflick), pp. 457–510. New York, NY: Van Nostrand Reinhold.

33 Reznick, D., Bryant, M. J. & Bashey, F. 2002 r- and K-selection revisited: the role of population regulation in life-history evolution. *Ecology* **83**, 1509–1520.

(doi:10.1890/0012-9658(2002)083[1509:RAKSRT]2.0.CO;2)

34 Brodmann, K. 1909 *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues.* Leipzig, Germany: Johann Ambrosius Barth. (English translation by Laurence Garey, 2006 *Brodmann's localisation in the cerebral cortex.* New York, NY: Springer.)

35 Geschwind, N. 1965 Disconnexion syndromes in animals and man. I. *Brain* **88**, 237–294. (doi:10.1093/brain/88.2.237)

36 Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G. & Orban, G. A. 2009 The representation of tool use in humans and monkeys: common and uniquely human features. *J. Neurosci.* **29**, 11 523–11 539. (doi:10.1523/JNEUROSCI.2040-09.2009)

37 Arbib, M. A. & Bota, M. 2006 Neural homologies and the grounding of neurolinguistics. In *Action to language via the mirror neuron system* (ed. M. A. Arbib), pp. 136–173. Cambridge, MA: Cambridge University Press.

38 Petrides, M. & Pandya, D. N. 2009 Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* **7**, e1000170. (doi:10.1371/journal.pbio.1000170)

39 von Bonin, G. & Bailey, P. 1974 *The neocortex of Macaca mulatta.* Urbana, IL: University of Illinois Press.

40 von Economo, C. 1929 *The cytoarchitectonics of the human cerebral cortex.* London, UK: Oxford University Press.

41 Jancke, L., Kleinschmidt, A., Mirzazade, S., Shah, N. J. & Freund, H. J. 2001 The role of the inferior parietal cortex in linking the tactile perception and manual construction of object shapes. *Cereb. Cortex* **11**, 114–121. (doi:10.1093/cercor/11.2.114)

42 Durand, J. B., Peeters, R., Norman, J. F., Todd, J. T. & Orban, G. A. 2009 Parietal regions processing visual 3D shape extracted from disparity. *Neuroimage* **46**, 1114–1126. (doi:10.1016/j.neuroimage.2009.03.023)

43 Jacob, S. N. & Nieder, A. 2009 Notation-independent representation of fractions in the human parietal cortex. *J. Neurosci.* **29**, 4652–4657. (doi:10.1523/JNEUROSCI.0651-09.2009)

44 Cappelletti, M., Lee, H. L., Freeman, E. D. & Price, C. J. 2010 The role of right and left parietal lobes in the conceptual processing of numbers. *J. Cogn. Neurosci.* **22**, 331–346. (doi:10.1162/jocn.2009.21246)

45 Castelli, F., Glaser, D. E. & Butterworth, B. 2006 Discrete and analogue quantity processing in the parietal lobe: a functional MRI study. *Proc. Natl Acad. Sci. USA* **103**, 4693–4698. (doi:10.1073/pnas.0600444103)

46 Tudusciuc, O. & Nieder, A. 2009 Contributions of primate prefrontal and posterior parietal cortices to length and numerosity representation. *J. Neurophysiol.* **101**, 2984–2994. (doi:10.1152/jn.90713.2008)

47 Knops, A., Thirion, B., Hubbard, E. M., Michel, V. & Dehaene, S. 2009 Recruitment of an area involved in eye movements during mental arithmetic. *Science* **324**, 1583–1585. (doi:10.1126/science.1171599)

48 Franklin, M. S. & Jonides, J. 2009 Order and magnitude share a common representation in parietal cortex. *J. Cogn. Neurosci.* **21**, 2114–2120. (doi:10.1162/jocn.2008.21181)

49 Malhotra, P., Coulthard, E. J. & Husain, M. 2009 Role of right posterior parietal cortex in maintaining attention to spatial locations over time. *Brain* **132**, 645–660. (doi:10.1093/brain/awn350)

50 Balan, P. F. & Gottlieb, J. 2009 Functional significance of nonspatial information in monkey lateral intraparietal area. *J. Neurosci.* **29**, 8166–8176. (doi:10.1523/JNEUROSCI.0243-09.2009)

51 Oristaglio, J., Schneider, D. M., Balan, P. F. & Gottlieb, J. 2006 Integration of visuospatial and effector information during symbolically cued limb movements in monkey lateral intraparietal area. *J. Neurosci.* **26**, 8310–8319. (doi:10.1523/JNEUROSCI.1779-06.2006)

52 Sereno, A. B. & Amador, S. C. 2006 Attention and memory-related responses of neurons in the lateral intraparietal area during spatial and shape-delayed match-to-sample tasks. *J. Neurophysiol.* **95**, 1078–1098. (doi:10.1152/jn.00431.2005)

53 Sereno, A. B. & Maunsell, J. H. 1998 Shape selectivity in primate lateral intraparietal cortex. *Nature* **395**, 500–503. (doi:10.1038/26752)

54 Gifford, G. W. & Cohen, Y. E. 2005 Spatial and non-spatial auditory processing in the lateral intraparietal area. *Exp. Brain Res.* **162**, 509–512. (doi:10.1007/s00221-005-2220-2)

55 Medendorp, W. P., Goltz, H. C. & Vilis, T. 2005 Remapping the remembered target location for anti-saccades in human posterior parietal cortex. *J. Neurophysiol.* **94**, 734–740. (doi:10.1152/jn.01331.2004)

56 Hu, S., Bu, Y., Song, Y., Zhen, Z. & Liu, J. 2009 Dissociation of attention and intention in human posterior parietal cortex: an fMRI study. *Eur. J. Neurosci.* **29**, 2083–2091. (doi:10.1111/j.1460-9568.2009.06757.x)

57 Zenon, A., Filali, N., Duhamel, J. R. & Olivier, E. 2010 Salience representation in the parietal and frontal cortex. *J. Cogn. Neurosci.* **22**, 918–930. (doi:10.1162/jocn.2009.21233)

58 Schendan, H. E., Amick, M. M. & Cronin-Golomb, A. 2009 Role of a lateralized parietal-basal ganglia circuit in hierarchical pattern perception: evidence from Parkinson's disease. *Behav. Neurosci.* **123**, 125–136. (doi:10.1037/a0013734)

59 Phillips, J. S., Velanova, K., Wolk, D. A. & Wheeler, M. E. 2009 Left posterior parietal cortex participates in both task preparation and episodic retrieval. *Neuroimage* **46**, 1209–1221. (doi:10.1016/j.neuroimage.2009.02.044)

60 McNab, F., Varrone, A., Farde, L., Jucaite, A., Bystritsky, P., Forssberg, H. & Klingberg, T. 2009 Changes in cortical dopamine D1 receptor binding associated with cognitive training. *Science* **323**, 800–802. (doi:10.1126/science.1166102)

61 Kawasaki, M., Watanabe, M., Okuda, J., Sakagami, M. & Aihara, K. 2008 Human posterior parietal cortex maintains color, shape and motion in visual short-term memory. *Brain Res.* **1213**, 91–97. (doi:10.1016/j.brainres.2008.03.037)

62 Rawley, J. B. & Constantinidis, C. 2009 Neural correlates of learning and working memory in the primate posterior parietal cortex. *Neurobiol. Learn Mem.* **91**, 129–138. (doi:10.1016/j.nlm.2008.12.006)

63 Oztekin, I., McElree, B., Staresina, B. P. & Davachi, L. 2009 Working memory retrieval: contributions of the left prefrontal cortex, the left posterior parietal cortex, and the hippocampus. *J. Cogn. Neurosci.* **21**, 581–593. (doi:10.1162/jocn.2008.21016)

64 Mitchell, D. J. & Cusack, R. 2008 Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. *Cereb. Cortex* **18**, 1788–1798. (doi:10.1093/cercor/bhm205)

65 Leon, M. I. & Shadlen, M. N. 2003 Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* **38**, 317–327. (doi:10.1016/S0896-6273(03)00185-5)

66 Bueti, D., Walsh, V., Frith, C. & Rees, G. 2008 Different brain circuits underlie motor and perceptual representations of temporal intervals. *J. Cogn. Neurosci.* **20**, 204–214. (doi:10.1162/jocn.2008.20017)

67 Coull, J. T., Vidal, F., Nazarian, B. & Macar, F. 2004 Functional anatomy of the attentional modulation of time estimation. *Science* **303**, 1506–1508. (doi:10.1126/science.1091573)

68 Davis, B., Christie, J. & Rorden, C. 2009 Temporal order judgments activate temporal parietal junction. *J. Neurosci.* **29**, 3182–3188. (doi:10.1523/JNEUROSCI.5793-08.2009)

69 Calton, J. L., Dickinson, A. R. & Snyder, L. H. 2002 Non-spatial, motor-specific activation in posterior parietal cortex. *Nat. Neurosci.* **5**, 580–588. (doi:10.1038/nn0602-862)

70 Beudel, M., Renken, R., Leenders, K. L. & de Jong, B. M. 2009 Cerebral representations of space and time. *Neuroimage* **44**, 1032–1040. (doi:10.1016/j.neuroimage.2008.09.028)

71 Oliveri, M., Koch, G., Salerno, S., Torriero, S., Lo Gerfo, E. & Caltagirone, C. 2009 Representation of time intervals in the right posterior parietal cortex: implications for a mental time line. *Neuroimage* **46**, 1173–1179. (doi:10.1016/j.neuroimage.2009.03.042)

72 Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)

73 Ishida, H., Nakajima, K., Inase, M. & Murata, A. 2010 Shared mapping of own and others' bodies in visuotactile bimodal area of monkey parietal cortex. *J. Cogn. Neurosci.* **22**, 83–96. (doi:10.1162/jocn.2009.21185)

74 Pelphrey, K. A., Morris, J. P. & McCarthy, G. 2004 Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* **16**, 1706–1716. (doi:10.1162/0898929042947900)

75 Muhlau, M. *et al.* 2005 Left inferior parietal dominance in gesture imitation: an fMRI study. *Neuropsychologia* **43**, 1086–1098. (doi:10.1016/j.neuropsychologia.2004.10.004)

76 Jackson, P. L., Meltzoff, A. N. & Decety, J. 2006 Neural circuits involved in imitation and perspective-taking. *Neuroimage* **31**, 429–439. (doi:10.1016/j.neuroimage.2005.11.026)

77 Dinstein, I., Gardner, J. L., Jazayeri, M. & Heeger, D. J. 2008 Executed and observed movements have different distributed representations in human aIPS. *J. Neurosci.* **28**, 11 231–11 239. (doi:10.1523/JNEUROSCI.3585-08.2008)

78 Yamazaki, Y., Yokochi, H., Tanaka, M., Okanoya, K. & Iriki, A. 2010 Potential role of monkey inferior parietal neurons coding action semantic equivalences as precursors of parts of speech. *Soc. Neurosci.* **5**, 105–117. (doi:10.1080/17470910802625306)

79 Culham, J. C. & Valyear, K. F. 2006 Human parietal cortex in action. *Curr. Opin. Neurobiol.* **16**, 205–212. (doi:10.1016/j.conb.2006.03.005)

80 Johnson-Frey, S. H., Newman-Norlund, R. & Grafton, S. T. 2005 A distributed left hemisphere network active during planning of everyday tool use skills. *Cereb. Cortex* **15**, 681–695. (doi:10.1093/cercor/bhh169)

81 Lewis, J. W. 2006 Cortical networks related to human use of tools. *Neuroscientist* **12**, 211–231. (doi:10.1177/1073858406288327)

82 Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., Onoe, H. & Iriki, A. 2001 Functional brain mapping of monkey tool use. *Neuroimage* **14**, 853–861. (doi:10.1006/nimg.2001.0878)

83 Agnew, Z. & Wise, R. J. 2008 Separate areas for mirror responses and agency within the parietal operculum.

J. Neurosci. **28**, 12 268–12 273. (doi:10.1523/JNEURO SCI.2836-08.2008)

84 Shmuelof, L. & Zohary, E. 2008 Mirror-image representation of action in the anterior parietal cortex. *Nat. Neurosci.* **11**, 1267–1269. (doi:10.1038/nn.2196)

85 Yamakawa, Y., Kanai, R., Matsumura, M. & Naito, E. 2009 Social distance evaluation in human parietal cortex. *PLoS ONE* **4**, e4360. (doi:10.1371/journal. pone.0004360)

86 Neggers, S. F., Van der Lubbe, R. H., Ramsey, N. F. & Postma, A. 2006 Interactions between ego- and allocentric neuronal representations of space. *Neuroimage* **31**, 320–331. (doi:10.1016/j.neuroimage.2005.12.028)

87 Decety, J., Jackson, P. L., Sommerville, J. A., Chaminade, T. & Meltzoff, A. N. 2004 The neural bases of cooperation and competition: an fMRI investigation. *Neuroimage* **23**, 744–751. (doi:10.1016/ j.neuroimage.2004.05.025)

88 Goldenberg, G. & Spatt, J. 2009 The neural basis of tool use. *Brain* **132**, 1645–1655. (doi:10.1093/brain/awp080)

89 Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J. & DeYoe, E. A. 2005 Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* **25**, 5148–5158. (doi:10.1523/JNEUROSCI.0419-05.2005)

90 Pelgrims, B., Andres, M. & Olivier, E. 2009 Double dissociation between motor and visual imagery in the posterior parietal cortex. *Cereb. Cortex* **19**, 2298–2307. (doi:10.1093/cercor/bhn248)

91 Schicke, T., Muckli, L., Beer, A. L., Wibral, M., Singer, W., Goebel, R., Rösler, F. & Röder, B. 2006 Tight covariation of BOLD signal changes and slow ERPs in the parietal cortex in a parametric spatial imagery task with haptic acquisition. *Eur. J. Neurosci.* **23**, 1910–1918. (doi:10.1111/j.1460-9568.2006.04720.x)

92 Zatorre, R. J., Halpern, A. R. & Bouffard, M. 2010 Mental reversal of imagined melodies: a role for the posterior parietal cortex. *J. Cogn. Neurosci.* **22**, 775–789. (doi:10.1162/jocn.2009.21239)

93 Champod, A. S. & Petrides, M. 2007 Dissociable roles of the posterior parietal and the prefrontal cortex in manipulation and monitoring processes. *Proc. Natl Acad. Sci. USA* **104**, 14 837–14 842. (doi:10.1073/ pnas.0607101104)

94 Kamigaki, T., Fukushima, T. & Miyashita, Y. 2009 Cognitive set reconfiguration signaled by macaque posterior parietal neurons. *Neuron* **61**, 941–951. (doi:10. 1016/j.neuron.2009.01.028)

95 Liston, C., Matalon, S., Hare, T. A., Davidson, M. C. & Casey, B. J. 2006 Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron* **50**, 643–653. (doi:10.1016/j.neuron.2006.04.015)

96 Nakahara, K., Hayashi, T., Konishi, S. & Miyashita, Y. 2002 Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science* **295**, 1532–1536. (doi:10.1126/science.1067653)

97 Pessoa, L., Rossi, A., Japee, S., Desimone, R. & Ungerleider, L. G. 2009 Attentional control during the transient updating of cue information. *Brain Res.* **1247**, 149–158. (doi:10.1016/j.brainres.2008.10.010)

98 Shomstein, S. & Yantis, S. 2006 Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. *J. Neurosci.* **26**, 435–439. (doi:10.1523/ JNEUROSCI.4408-05.2006)

99 Stoet, G. & Snyder, L. H. 2004 Single neurons in posterior parietal cortex of monkeys encode cognitive set. *Neuron* **42**, 1003–1012. (doi:10.1016/j.neuron. 2004.06.003)

100 Ogawa, T. & Komatsu, H. 2009 Condition-dependent and condition-independent target selection in the macaque posterior parietal cortex. *J. Neurophysiol.* **101**, 721–736. (doi:10.1152/jn.90817.2008)

101 Hampshire, A., Gruszka, A., Fallon, S. J. & Owen, A. M. 2008 Inefficiency in self-organized attentional switching in the normal aging population is associated with decreased activity in the ventrolateral prefrontal cortex. *J. Cogn. Neurosci.* **20**, 1670–1686. (doi:10.1162/jocn. 2008.20115)

102 Crowe, D. A., Chafee, M. V., Averbeck, B. B. & Georgopoulos, A. P. 2004 Participation of primary motor cortical neurons in a distributed network during maze solution: representation of spatial parameters and time-course comparison with parietal area 7a. *Exp. Brain Res.* **158**, 28–34. (doi:10.1007/ s00221-004-1876-3)

103 Freedman, D. J. & Assad, J. A. 2009 Distinct encoding of spatial and nonspatial visual information in parietal cortex. *J. Neurosci.* **29**, 5671–5680. (doi:10.1523/ JNEUROSCI.2878-08.2009)

104 Hoenig, K. & Scheef, L. 2009 Neural correlates of semantic ambiguity processing during context verification. *Neuroimage* **45**, 1009–1019. (doi:10.1016/j.neuro image.2008.12.044)

105 Yang, T. & Shadlen, M. N. 2007 Probabilistic reasoning by neurons. *Nature* **447**, 1075–1080. (doi:10.1038/ nature05852)

106 Tosoni, A., Galati, G., Romani, G. L. & Corbetta, M. 2008 Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. *Nat. Neurosci.* **11**, 1446–1453. (doi:10.1038/nn.2221)

107 Turriziani, P., Oliveri, M., Bonni, S., Koch, G., Smirni, D. & Cipolotti, L. 2009 Exploring the relationship between semantics and space. *PLoS ONE* **4**, e5319. (doi:10.1371/journal.pone.0005319)

108 Yamazaki, Y., Hashimoto, T. & Iriki, A. 2009 The posterior parietal cortex and non-spatial cognition. *F1000 Biol. Rep.* **1**, 74. (doi:10.3410/B1-74)

109 Arbib, M. A. & Lee, J. 2008 Describing visual scenes: towards a neurolinguistics based on construction grammar. *Brain Res.* **1225**, 146–162. (doi:10.1016/j.brain res.2008.04.075)

110 Almor, A., Smith, D. V., Bonilha, L., Fridriksson, J. & Rorden, C. 2007 What is in a name? Spatial brain circuits are used to track discourse references. *Neuroreport* **18**, 1215–1219. (doi:10.1097/WNR.0b013e 32810 f2e11)

111 Ogawa, A., Yamazaki, Y., Ueno, K., Cheng, K. & Iriki, A. 2010 Neural correlates of species-typical illogical cognitive bias in human inference. *J. Cogn. Neurosci.* **22**, 2120–2130. (doi:10.1162/jocn.2009.21330)

112 Ogawa, A., Yamazaki, Y., Ueno, K., Cheng, K. & Iriki, A. 2010 Inferential reasoning by exclusion recruits parietal and prefrontal cortices. *Neuroimage* **52**, 1603–1610. (doi:10.1016/j.neuroimage.2010.05.040)

113 Hihara, S., Yamada, H., Iriki, A. & Okanoya, K. 2003 Spontaneous vocal differentiation of coo-calls for tools and food in Japanese monkeys. *Neurosci. Res. (NY)* **45**, 383–389.

114 Deacon, T. 1997 *The symbolic species.* London, UK: Penguin Press.

115 Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. 2005 *From DNA to diversity: molecular genetics and the evolution of animal design*, 2nd edn. Malden, MA: Blackwell Publishing.

116 De Robertis, E. M. & Sasai, Y. 1996 A common plan for dorsoventral patterning in Bilateria. *Nature* **380**, 37–40. (doi:10.1038/380037a0)

117 Mizunami, M., Yokohari, F. & Takahata, M. 1999 Exploration into the adaptive design of the arthropod 'microbrain'. *Zool. Sci.* **16**, 703–709. (doi:10.2108/zsj.16.703)

118 Mizunami, M., Yokohari, F. & Takahata, M. 2004 Further exploration into the adaptive design of the arthropod 'microbrain': I. Sensory and memory-processing systems. *Zool. Sci.* **21**, 1141–1151. (doi:10.2108/zsj.21.1141)

119 Ward, A. B., Weigl, P. D. & Conroy, R. M. 2002 Functional morphology of raptor hindlimbs: implications for resource partitioning. *Auk* **119**, 1052–1063. (doi:10.1642/0004-8038(2002)119[1052:FMORHI]2.0.CO;2)

120 Martin, G. R. & Katzir, G. 1994 Visual fields and eye movements in herons (Ardeidae). *Brain Behav. Evol.* **44**, 74–85. (doi:10.1159/000113571)

121 White, C. R., Day, N., Butler, P. J. & Martin, G. R. 2007 Vision and foraging in cormorants: more like herons than hawks? *PLoS ONE* **2**, e639. (doi:10.1371/journal.pone.0000639)

122 Ankel-Simons, F. 2007 *Primate anatomy: an introduction*, 3rd edn. Burlington, VT: Academic Press.

123 Larson, S. G. & Stern Jr, J. T. 2006 Maintenance of above-branch balance during primate arboreal quadrupedalism: coordinated use of forearm rotators and tail motion. *Am. J. Phys. Anthropol.* **129**, 71–81. (doi:10.1002/ajpa.20236)

124 Marzke, M. W. 1996 Evolution of the hand and the bipedality. In *Handbook of human symbolic evolution* (eds A. Lock & C. R. Peters), pp. 126–154. Oxford, UK: Oxford University Press.

125 Buneo, C. A. & Andersen, R. A. 2006 The posterior parietal cortex: sensorimotor interface for the planning and online control of visually guided movements. *Neuropsychologia* **44**, 2594–2606. (doi:10.1016/j.neuropsychologia.2005.10.011)

Research

Individual and social learning processes involved in the acquisition and generalization of tool use in macaques

S. Macellini[1,†], M. Maranesi[2,†], L. Bonini[2], L. Simone[2], S. Rozzi[2], P. F. Ferrari[1,2,*] and L. Fogassi[2,3,*]

[1]*Dipartimento di Biologia Evolutiva e Funzionale, Università di Parma, via Usberti 11/a, 43124 Parma, Italy*
[2]*Dipartimento di Neuroscienze and Istituto Italiano di Tecnologia (IIT), Brain Center for Motor and Social Cognition, Parma, via Volturno 39, 43125 Parma, Italy*
[3]*Dipartimento di Psicologia, Università di Parma, Borgo Carissimi 10, 43121 Parma, Italy*

Macaques can efficiently use several tools, but their capacity to discriminate the relevant physical features of a tool and the social factors contributing to their acquisition are still poorly explored. In a series of studies, we investigated macaques' ability to generalize the use of a stick as a tool to new objects having different physical features (study 1), or to new contexts, requiring them to adapt the previously learned motor strategy (study 2). We then assessed whether the observation of a skilled model might facilitate tool-use learning by naive observer monkeys (study 3). Results of study 1 and study 2 showed that monkeys trained to use a tool generalize this ability to tools of different shape and length, and learn to adapt their motor strategy to a new task. Study 3 demonstrated that observing a skilled model increases the observers' manipulations of a stick, thus facilitating the individual discovery of the relevant properties of this object as a tool. These findings support the view that in macaques, the motor system can be modified through tool use and that it has a limited capacity to adjust the learnt motor skills to a new context. Social factors, although important to facilitate the interaction with tools, are not crucial for tool-use learning.

Keywords: body schema; tool selection; sensorimotor experience; action perception; mirror neurons

1. INTRODUCTION

It is well acknowledged that several primate species are capable of selecting and using tools. In chimpanzees and capuchin monkeys, the flexible and the extensive use of tools have been widely documented both in free-living [1–3] and captive populations [4,5]. More scattered and limited are reports on tool use in free-ranging populations of macaques. Most of these studies documented the use of tools only in a limited number of individuals and often the reports are anecdotal [6]. More recently, it has been reported that long-tailed macaques of a wild population in Thailand regularly use stones as tools to crack shelled seafood [7]. Despite this example, there is still a general agreement in the scientific community that macaques are not skilled tool users. However, this picture becomes more complex, if one considers studies on macaques

under more controlled experimental conditions or in captivity, where long-lasting observations are feasible.

In fact, more detailed descriptions are available on captive and free-ranging provisioned groups of macaques, in which prolonged observations allowed researchers to understand which factors facilitate the acquisition of tool use or prevent individuals from acquiring new behaviours [8–13].

A series of laboratory experiments demonstrated that macaques are capable of learning the use of tools for retrieving food out of reach [14,15]. Under certain circumstances, the process of tool-use learning may require a relatively short time of training through instrumental conditioning procedures. Other laboratory studies demonstrated that macaques can learn to use even more complex tools, such as pliers, requiring a higher level of coordination and handedness [16]. Together, these studies indicate that macaques are capable of refining their motor representations and have the cognitive potential to include the tools within their expanded motor repertoire. This is also supported by neurophysiological studies indicating that, after tool-use learning, motor representations and body schema change in the parietal and premotor cortices [16,17].

An important issue that has been very scarcely investigated is how can monkeys discriminate the appropriate tool for a given task. Recent work in capuchin monkeys showed that they can select the most adequate tool for extracting food protected in a nut shell based on functional features such as weight and shape [2]. Few studies have explored in detail how non-human animals represent tools and, in particular, whether they distinguish between functional and non-functional objects based on their physical features [14,15,18]. In a series of experiments on object knowledge, Hauser [19] assessed in cotton-top tamarins the capacity of understanding which properties of a tool are relevant to its functioning. Once monkeys learned to use canes to retrieve food out of reach, they were presented with a variety of new tools with different colours, shapes, textures and sizes. The results showed that, on average, monkeys chose the functional tool. This has been interpreted as a demonstration that they use a strategy based on an understanding of the means-end relationship [19]. However, it cannot be excluded that trial-and-error processes, based on the sensorimotor experience with the tool, could have intervened. This latter interpretation would also be consistent with the conclusions reached by Visalberghi & Limongelli [20], based on their studies on capuchin monkeys.

Considering the reports so far reviewed about tool use in macaques in the wild, captivity and in more controlled experimental conditions [6,8–15], the issue of tool selection and of the underlying cognitive processes remains still poorly investigated in this taxon.

The process of tool-use acquisition can also be facilitated by social factors. Several studies showed in different species of macaques the possibility to acquire new behaviours through social-based learning. This social transmission of tool use can account for traditions and cultures so well developed and documented in primates [21].

One of the possibilities for a naive subject to learn a new behaviour is to observe an expert individual performing the action. The acquisition of the new behaviour will be probably linked also to the frequency with which the subject observes the demonstrator performing that action [22]. Other factors that increase the likelihood of social learning new behaviours are the attention that the subject pays towards the observed behaviour and its proximity to the demonstrator [22–30]. Thus, the acquisition of a new behaviour in naive subjects should be faster in those individuals with greater opportunity to observe and learn from expert models [22,26,27]. This observer–demonstrator paradigm has been typically used in captivity and in laboratory settings [21].

The series of experiments we present here had two main objectives and were organized into two main parts.

The first part was focused on the issue of individual learning processes. In particular, it was aimed at investigating whether, and to what extent, monkeys previously trained to use a tool for retrieving food could generalize their capacity across different tools with novel features and different contexts requiring the adjustment of the learned motor strategy.

In the second part, the monkeys that were employed for individual learning experiments during the first part,

served as demonstrators of tool use for completely naive macaque monkeys. In this observer–demonstrator paradigm, we explored the possible presence of social learning processes in the observing individuals.

2. STUDY 1: SELECTION OF TOOLS BASED ON THEIR PHYSICAL PROPERTIES

Capuchin monkeys and chimpanzees are capable of selecting tools with different features in relation to their behavioural purposes [2,5]. However, this ability is still largely unexplored in macaques. The main aims of this study were twofold: (i) to assess how the learned capacity to use a tool in a specific task can be generalized to other types of tools having the same length, and (ii) to verify whether monkeys can select, among tools of different shapes, one of appropriate length to enable food retrieval.

Monkeys were first trained to use a stick in order to retrieve a creamy food (yogurt) out of arm's reach. Then, they were presented with two novel elongated tools of different shape, in addition to the stick. In one condition, all tools were functional for retrieving food, in another condition only one was functional, the other two being too short for this purpose.

We analysed the animals' choice in both conditions.

(a) *Material and methods*
(i) *Subjects*
The experimental subjects were two male pigtailed macaque monkeys (*Macaca nemestrina*), here identified as M1 and M2. Both monkeys were captive born, mother-reared until they were 2–3 years old and then individually housed at the Primate Section of the Department of Neuroscience, University of Parma. At the time of testing, they were both 5 years old. Both monkeys were singly housed in cages ($175 \times 100 \times 100$ cm) allowing them visual and auditory contact with other monkeys (*Macaca nemestrina* and *Macaca mulatta*) housed in the same room.

To maintain a high motivation to the task, during experiments, subjects were mildly food-deprived, receiving their daily food only at the end of each testing session. Food consisted of fresh fruits, vegetables, bread, seeds and monkey chow. Water was always available.

(ii) *Apparatus and training procedures*
The basic set-up employed in this study is shown in figure 1a.

During the first phase, monkeys sat in their home cages. They were allowed to retrieve the food (yogurt) from a transparent Plexiglas container (inner diameter 6.5 cm, height 5.5 cm) by means of a wooden stick (diameter 1 cm, length 22 cm). The container was located in front of the monkey cage, screwed on a plywood table (length 70 cm, width 75 cm, height 32 cm from the floor of the cage) outside the monkey reaching space (44.5 cm from the cage bars). In this phase, none of the monkeys spontaneously succeeded or attempted to use the tool for retrieving the food.

In the second phase, we tried to facilitate the monkeys to individually learn tool use. The experimenter inserted the stick into the container filled with yogurt, and then monkeys were allowed to retrieve the tool
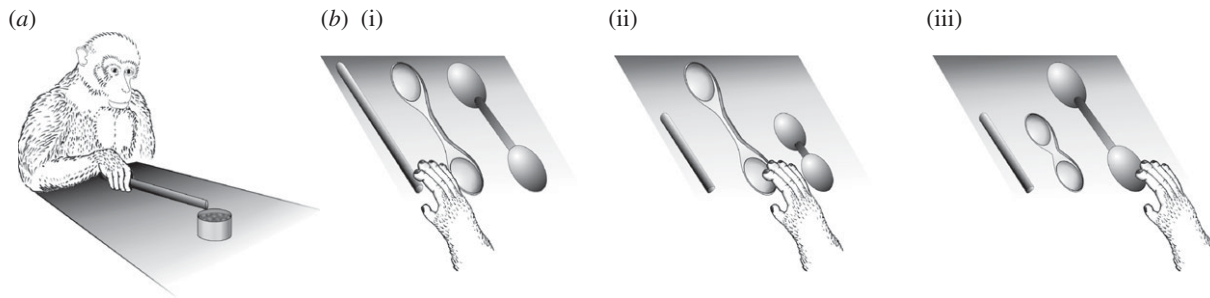
Figure 1. (*a*) Basic experimental set-up employed in study 1. (*b*) Experimental conditions and features of the tools used for testing tool selection during study 1: *b*(i) all three tools functional (3F); *b*(ii) functional spoon (FS) and non-functional stick and egg-shaped tools; *b*(iii) functional egg-shaped tool (FE) and non-functional stick and spoon.

and lick the food from it (10 sessions lasting 10 min, 10 trials per session). The intent underlying this procedure was that of prompting part of the motor sequences that the monkeys had to perform to accomplish the task (e.g. grasping the tool already inserted into the glass, bringing it to the mouth and eating the yogurt). Even though monkeys easily succeeded in retrieving the tool from the container and eating the yogurt in all trials, subsequently they did not show any attempt at spontaneous tool use. This rendered it necessary to introduce a shaping procedure. In the first part of this procedure, monkeys were first reinforced with food whenever they touched and grasped the tool. Then, they received food whenever they extended the arm while holding the tool and, finally, whenever they touched the container with it. This training procedure was employed twice a day for four consecutive weeks, until the monkeys successfully performed the correct action sequence with at least 90 per cent correct trials per session, for at least three consecutive sessions.

The training was recorded with a digital camcorder CANON MVX250i, and the video clips were subsequently analysed to evaluate the rate of success in the task.

(iii) *Experimental task*

Monkeys were tested in their home cage. In each session (5 min long), three wooden tools differing in shape (spoon, egg-shaped and stick; figure 1*b*), but not in texture and colour, were used. Each tool could be presented in one of two versions, either 'functional' or 'non- functional'. The functional tools were 22 cm long and enabled the monkey to reach for the food, while the non-functional ones were only 11 cm long, thus not long enough to reach it. The three tools were simultaneously presented to the monkey on the same plywood table previously used during the training phase, in the following combinations:

— all three functional (3F);
— functional spoon (FS), non-functional stick and egg-shaped tool; and
— functional egg-shaped tool (FE), non-functional stick and spoon.

Each combination was presented three times, resulting in a total number of nine sessions for each animal. In each session, monkeys were free to interact with any of the available tools and try to use each of them for reaching the food. The order of presentation was as follows: 3F, FS, FE, FS, FE, 3F, FE, 3F, FS.

There was no session in which the only functional tool was the stick, because this was the most familiar to the animal and could have biased its choice.

(iv) *Behavioural analysis*

All sessions were video recorded with a digital camcorder CANON MVX250i and the tapes independently analysed by two experimenters familiar with the experimental procedure. The frequency of interactions of the monkey with each of the available tools in each session was assessed. An interaction was defined as the grasping of a tool followed by the attempt to insert it into the container, regardless of the outcome of the attempt (the rate of success was always above 90%). For each session, we also scored which tool was the first contacted by the monkey, in order to verify whether its choice was based on an evaluation of the suitability of the physical features of the tool in relation to task requirements, or on a mere trial-and-error learning process.

(v) *Statistical analysis*

χ^2-tests were applied to assess whether there was a preference for a specific type of tool during the 3F sessions. The same test was then employed in the sessions in which only one tool was functional (i.e. FS and FE) to assess whether the general choice frequency for the functional tool was higher than that expected based on chance. Furthermore, in these sessions, we also tested whether the frequency of choice of the functional tool as first was higher than chance.

(b) *Results*

The results of study 1 are summarized in figure 2. In the sessions in which all tools were functional (i.e. 3F), both monkeys preferred the stick to retrieve food significantly above chance level (M1 $\chi^2 = 218.49$, $p < 0.001$ and M2 $\chi^2 = 16.95$, $p < 0.001$). During the sessions in which only one tool was functional (FS and FE), both monkeys more frequently used the appropriate tool (M1 $\chi^2 = 115.996$, $p < 0.001$ and M2 $\chi^2 = 8.00$, $p < 0.005$). More specifically, both M1 and M2 used the egg-shaped tool in the FE sessions ($\chi^2 = 43.29$, $p < 0.001$ and $\chi^2 = 8.40$, $p < 0.005$, respectively) more frequently than expected by chance. As far as the spoon is concerned, both monkeys chose it more frequently during FS sessions, although M1 did it at a higher frequency with respect to chance level ($\chi^2 = 74.67$, $p < 0.001$), while M2 did not reach a statistically significant level ($\chi^2 = 0.76$, n.s.).
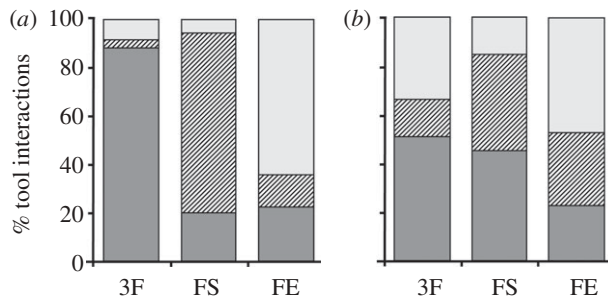
Figure 2. Percentage of tool interactions with each of the available tools. Each histogram represents the interactions during the sessions with all tool types: functional (3F), functional spoon (FS) and functional egg-shaped tool (FE) for (*a*) M1 and (*b*) M2. Light grey bars, egg-shaped; striped bars, spoon; dark grey bars, stick.

In the sessions in which only one tool was functional, the first-grasped tool was randomly chosen by both monkeys, contradicting the hypothesis that the functional tool was mostly preferred as a 'first choice' (M1 $\chi^2 = 0.75$, n.s. and M2 $\chi^2 = 0.83$, n.s.).

It is worth noting that, in all conditions, during the first interaction with an unfamiliar tool, the two monkeys never brought it to the mouth, but they used it to retrieve the food by applying the same motor patterns employed with the familiar tool. In only a few occasions, we did record exploratory behaviours consisting of tool manipulation without using it.

3. STUDY 2: GENERALIZATION OF TOOL-USE STRATEGY TO DIFFERENT ENVIRONMENTAL CONTEXTS

If tool-use skill acquired in a certain situation is based on a causal understanding of the physical properties of the object used as a tool, this capacity should be easily transferred to contexts different from that in which learning occurred. This transfer, here referred to as generalization, may be expressed through newly adapted behavioural strategies with various levels of complexity, depending on the contextual demands. The aim of this study was to assess whether monkeys capable of using a stick to extract food from a container (study 1) could generalize the learned skills to a new contextual setting.

More specifically, a transparent cylindrical container was located inside the monkeys' home cage, firmly fixed to the cage bars. Monkeys could directly interact with the container, but the small aperture on the top of it and its elongated shape allowed them to retrieve the food only by inserting the stick into the container (figure 3). Therefore, when compared with study 1, monkeys could adopt a wider range of motor strategies to retrieve the food, but it was designed so that the previously learned one was not effective in this context. Furthermore, we verified whether the presence/absence of food inside the container, which represents the monkey's behavioural goal, could affect its attempts to reach for the food.

(a) *Material and methods*
(i) *Subjects*
The experimental subjects were the same as in study 1.
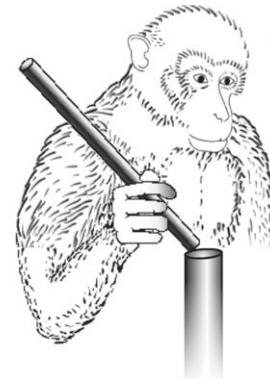
Figure 3. Schematic set-up and apparatus employed in study 2.

(ii) *Apparatus and procedures*
The food (yogurt) was put into a cylindrical Plexiglas container (internal diameter 3 cm, height 18 cm) positioned inside the cage and firmly attached to the bars. At the beginning of each trial, the tool, a stick identical to that used during study 1, was placed on the floor of the cage. The monkey had to grasp the stick and to insert it completely inside the container to obtain the food.

The study was divided into four phases: *Baseline*, *Familiarization/Facilitation*, *Practice* and *Food/No food test*. During Baseline (six sessions, each 5 min long), we evaluated the frequency of interactions between the monkey's hand and the tool. Since none of the monkeys succeeded in getting food by using the tool, we introduced the Familiarization/Facilitation phase. This phase consisted of a single experimental session, lasting 40 min, during which the experimenter inserted the stick twice into the container, so that the monkey could simply retrieve the stick and lick the yogurt. We assessed the monkeys' behaviour scoring both their failed attempts and autonomously performed trials of successful tool use during the whole session. The Practice phase consisted of three sessions, each 10 min long, in which the monkey was given the tool to accomplish autonomously the same task. The fourth phase, Food/No food test, consisted of six sessions. In three of them, the container was filled with yogurt (*Food sessions*, F), while in the other three it was empty (*No food sessions*, NF). Before using it for NF sessions, the container was thoroughly washed to remove residual food or smells. Sessions F and NF were alternated and lasted 15 min each: during the first 5 min of each session, the monkeys were given the possibility to explore the container, filled or empty, allowing them to familiarize with the container in the absence of the tool. Subsequently, the experimenter introduced the tool in the cage and the monkey was allowed to use it for the next 10 min of the session.

(iii) *Behavioural analysis*
All the experimental phases were recorded with a digital camcorder CANON MVX250i and the videos were off-line independently analysed by two coders familiar with the experimental phases. The frequency of interactions of the monkeys with the tool was scored. Since it was difficult to systematically describe and categorize the numerous patterns of behaviour each monkey displayed with the tool, we limited our analysis to a few items relative to those events in which the
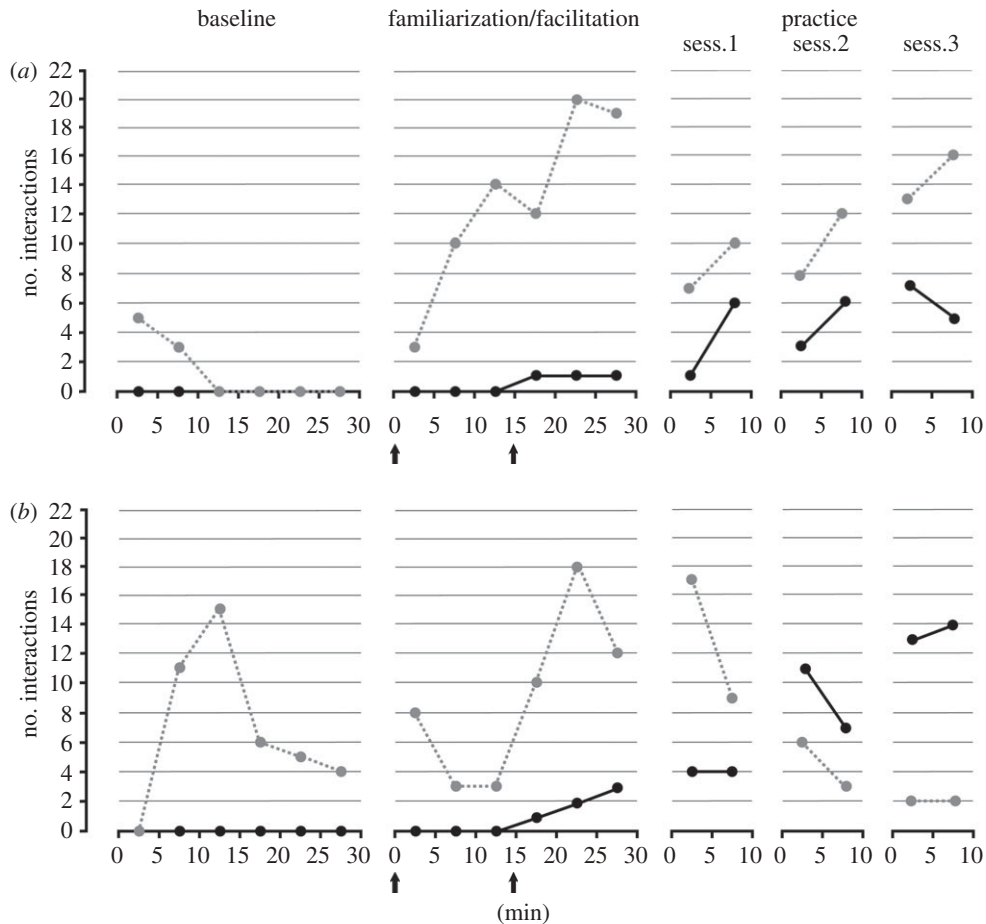
Figure 4. Number of tool insertions and insertion attempts performed by (*a*) M1 and (*b*) M2 during the Baseline, Familiarization/Facilitation and Practice phases. The arrows indicate the time points when the experimenter inserted the stick inside the container, providing the monkey with a cue about part of the motor sequence to be done. Grey circles with dashed line, attempt; black circles with continuous line, insertion.

monkey grasped the tool and directed it towards the aperture on top of the container, trying to insert it although unsuccessfully ('attempt'). When an interaction ended with the tool completely inserted into the container in F sessions, enabling food retrieval, it was considered as 'insertion'.

(iv) *Statistical analysis*

χ^2-tests were used to compare the number of interactions with the tool and insertions during F and NF phases, against those expected by chance.

(**b**) *Results*

Figure 4 shows the time course of the frequency of interactions with the tool (in terms of insertion and attempt, separately) of both monkeys across subsequent sessions and phases. During the Baseline phase, none of the subjects succeeded in retrieving the food with the tool. From the beginning of this phase, both monkeys tried to directly access the food by biting the container or probing their fingers inside it. After a few minutes in which they failed to reach the food with this strategy, both of them grasped the stick and brought it in proximity to the container. More specifically, M2 made several attempts to insert the tool into the container, but always without success. The tool was sometimes handled by the

monkey with its arms outside of the cage bars and lifted above the container, attempting to insert it. In other cases, the monkey manipulated the tool very vigorously with both hands in proximity to the container opening. Such episodes of tool manipulation were usually highly variable, jerky and poorly coordinated, so that in several cases, the tool fell on the floor. All of these behaviours decreased over time. M1 also approached the container with the stick and manipulated the tool in its proximity. However, when compared with M2, it performed only a few clear attempts of insertion, and its manipulations of the tool rapidly decreased over time.

During the Familiarization/Facilitation phase, following the first tool insertion by the experimenter, both monkeys rapidly increased their rate of attempts to use the tool for retrieving food (M1 $\chi^2 = 49.44$, $p < 0.001$; M2 $\chi^2 = 6.00$, $p < 0.05$). In the first minutes after stick extraction, both monkeys licked the tool and tried to lick the spilled drops of yogurt near the opening of the container or to probe with the fingers into the container to extract the left-over drops. These activities lasted a few minutes before monkeys started to use the tool again.

Interestingly, among other attempts, M1 tried on one occasion to replicate the exact motor pattern that was effective in study 1, namely, sitting in a
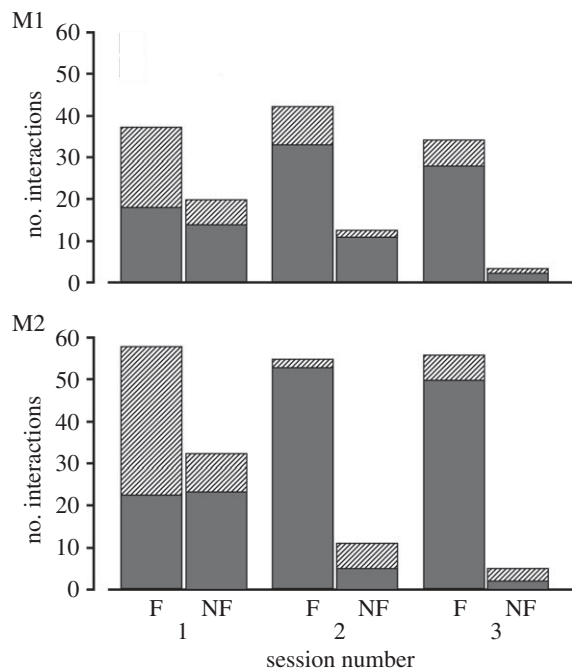
Figure 5. Number of tool insertions and attempts performed by M1 and M2 during the sessions in which the container was alternatively filled with yogurt (Food sessions, F) or empty (No Food sessions, NF). Striped bars, attempts; grey bars, insertions.

cage sector far from the container and then extending the arm holding the stick, pointing with it and attending to the container. Furthermore, both monkeys started to show some successful insertions of the tool into the container. In particular, M1 succeeded in retrieving food autonomously for the first time after 30 insertion attempts, 20 min after the beginning of the Familiarization/Facilitation phase, while M2 succeeded for the first time after 15 insertion attempts, 20 min after the beginning of the same phase.

During the Practice phase, the monkeys' rate of success tended to increase across sessions, but the difference between the first and the last Practice session was significant only in M2 (M1 $\chi^2 = 1.80$, n.s.; M2 $\chi^2 = 9.52$, $p < 0.01$).

In the Food/No food phase (figure 5), the number of interactions with the tool, calculated by pooling insertions and attempts frequencies, was significantly higher in the F than in NF condition (M1 $\chi^2 = 39.02$, $p < 0.001$; M2 $\chi^2 = 65.06$, $p < 0.001$). Nevertheless, both monkeys also tried in a consistent number of cases to insert the stick into the container in the NF condition, succeeding on several occasions in reaching the bottom of the container with the tip of the stick. Interestingly, after tool insertion in the empty container, both monkeys often extracted the stick and brought it to the mouth, smelling and licking it, thus behaving similarly to in the F condition. However, the frequency of interactions with the tool (taking together insertions and attempts frequencies) remained similar across sessions in the F condition (M1 $\chi^2 = 0.84$, n.s.; M2 $\chi^2 = 0.14$, n.s.), while it decreased significantly across sessions in the NF condition (M1 $\chi^2 = 12.17$, $p < 0.005$; M2 $\chi^2 = 25.12$, $p < 0.001$).

4. STUDY 3: EFFECT OF SOCIAL FACILITATION ON TOOL USE

Previous studies have shown that macaques can sometimes benefit from observing the behaviour of skilled individuals for learning novel actions [6,31]. However, it is still unclear whether the observation of a skilled conspecific can facilitate tool-use learning in an observer monkey, and what information macaques can extract by observing other individuals using a tool [21].

The aim of the present study was to assess the behavioural responses of naive macaque monkeys after exposure to a trained conspecific using a rake for retrieving a piece of food out of arm's reach. We scored the observers' behaviour both *after* they were exposed to the performing model and *during* the model's performance.

For this study, we used a rake as tool because it requires minimal handling abilities and effort, and it has to be moved mainly in a two-dimensional space in order to achieve the goal (i.e. retrieving the food placed on the table in front of the monkey).

(a) *Material and methods*
(i) *Subjects*
The subjects were nine male *Macaca mulatta*, aged from 5 to 7 years. They were all naive to the use of tools. The two *Macaca nemestrina* employed in study 1 and study 2 acted as models. The procedures to train the model monkeys to use the rake were similar to those described by Iriki and co-workers [32,33].

The rearing and housing conditions of the animals were the same as described in study 1 and study 2. The use of animals of the same genus (*Macaca*), but of different species (*M. mulatta*) with respect to the model (*M. nemestrina*) as observers should not represent theoretical problems, given the similarity in the general body morphology as well as in the patterns of communicative and non-communicative behaviours shared by the two species [34]. Finally, our monkeys were very familiar with each other and the nature of the task did not require any species-specific behaviour.

(ii) *Apparatus*
The handle of the rake was a wooden stick of 1 cm diameter and 35 cm length, while the head was a wooden splint (12.5 × 4.5 cm), fixed to the handle in its centre.

The model and the observer were in different cages, one facing the other and separated by plywood working tables on which the tool and the food were placed. We used the same type of working table described in the first study. The experimenter placed small pieces of food (a piece of apple) on the table top by introducing them from the bottom of the table, unseen by the monkeys, through holes (diameter 1.5 cm each). This procedure was employed in order to prevent the monkeys from being distracted by the experimenter's action and to reduce interference with their behaviour. One hole was made on the side of the table where the demonstrator was located, 57 cm outside its cage, so that it could be reached only by using the rake. A second hole was made on the same side, but at 34 cm from the cage, so that the monkey
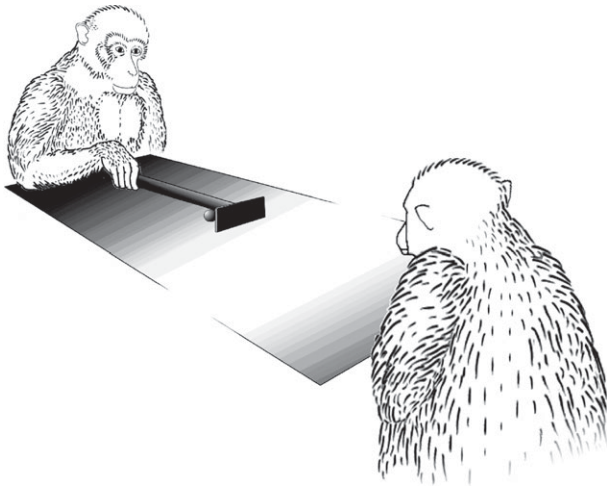
Figure 6. Set-up and apparatus employed in study 3. The picture shows the model retrieving a small piece of food using the rake while the observer is looking at the action.

could retrieve the food from it by hand. A third hole was made on the side of the table where the observer was located, 57 cm outside its cage.

(iii) *Tasks and procedure*

The experimental setting is shown in figure 6. Each trial started with the experimenter placing a piece of food on the table, on the side of the model monkey, through the hole located at 57 cm from the model's cage, so that the food was completely out of reach. Thus, the model monkey had to use the rake to reach the food and drag it along the table until it was graspable with the hand. In each phase, the model received a piece of apple every 15 s.

The experiment was divided into four phases: *Baseline*, *Observation-Delayed tool interaction*, *Observation-Simultaneous tool interaction* and *Follow-up*.

During Baseline (10 sessions, 10 min each), the model was given pieces of food by the experimenter through the closest hole present on its side of the table. While the model was engaged in grasping with the hand and eating the food, the observers were allowed to interact with the tool in order to retrieve a piece of food placed by the experimenter on the observer's side of the table, out of arm's reach.

The Observation-Delayed tool interaction phase consisted of 10 experimental sessions (25 min each). First, the model performed the correct food retrieving behaviour with the rake 40 times within a 10 min period (one trial every 15 s). Then, after a 5 min break, the observer was given a tool identical to that of the model for the following 10 min. The observer thus had the possibility to easily reach and grasp the rake in order to retrieve a piece of food placed on the table out of arm's reach. During this 10 min period, the model was given pieces of food by the experimenter through the hole closest to its cage, so that it could reach for and grasp it with the hand.

During Observation-Simultaneous tool interaction (10 sessions, 10 min each), the model and the observer were simultaneously provided with a rake and they could use it to reach a piece of food introduced by the experimenter every minute through the two farthest holes, on each side of the table.
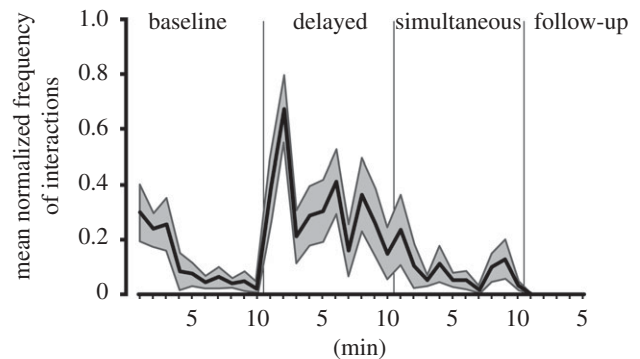
Figure 7. Mean normalized frequency of hand–tool interactions of all the observing subjects during the Baseline, Observation-Delayed tool interaction, Observation-Simultaneous tool interaction and Follow-up phases. Grey shaded area represents 1 s.e.

During the Follow-up phase (five sessions, 10 min each), the model was not present and the observer monkeys alone could interact with their own rake for the duration of the whole experimental session. This phase was aimed at assessing whether the observation of a conspecific using a tool in the previous phases affected the number of tool interactions by the observers in the absence of any model.

(iv) *Behavioural analysis*

All sessions were video recorded with a digital camcorder CANON MVX250i and the tapes analysed independently by two experimenters not blind to the experimental phases. A third scorer blind to the experimental conditions analysed 20 per cent of the sessions, showing a high concordance with the scores attributed by non-blind experimenters (Kendall $\tau = 0.89$, $p < 0.001$). The number of interactions of the monkeys' hands with the tool (touching or grasping) and the monkeys' attempts to use the tool for retrieving food were scored.

(v) *Statistical analysis*

The frequency of hand–tool interaction for each subject in each session has been normalized by dividing each value by the higher value recorded for that subject among the sessions of each specific phase. By this procedure, we obtained values ranging from 0 to 1, which allowed us to pool data of all tested subjects. A one-way repeated measure ANOVA, eventually followed by Newman–Keuls *post hoc* tests, was used in order to compare the number of hand–tool interactions of the observing monkeys among subsequent experimental phases.

(b) *Results*

None of the observing subjects tried to use the tool to retrieve food. However, the comparison of the number of hand–tool interactions revealed a significant difference among conditions ($F = 15.109$, $p < 0.001$). The *post hoc* test showed that only the interactions performed during the Observation-Delayed tool interaction were significantly more frequent with respect to those in all other experimental phases ($p < 0.01$; figure 7).

5. DISCUSSION

(a) *Individual learning and understanding of the physical properties of a tool*

The results of the first study demonstrate that monkeys in the current experimental setting were not capable of using a tool for retrieving food. Nevertheless, they could easily do so after a shaping procedure. Subsequently, they also showed the capacity to generalize the previously learned motor pattern to the use of unfamiliar tools. Moreover, the monkeys demonstrated they were capable of selecting a functional tool based on its length, despite its shape not being familiar to them. However, this capacity does not appear to be based on an understanding of the functional properties of the tool, but had to be achieved through a fast trial-and-error learning process.

Previous captive studies on macaques have reported that the spontaneous use of tools sometimes occurs [9,10], but the relatively natural setting of these experiments renders it difficult to identify which type of learning process leads to the result. Other studies showed that, under more controlled experimental conditions, macaques can learn to use tools without much practice or sensorimotor experience [14]. Surprisingly, some of the monkeys in this latter study succeeded in using the tool after only a few minutes of interaction with it, promoting the authors to interpret this finding as a sign of insightful behaviour. However, the macaques were partially restrained in a primate chair, so that the number of relevant stimuli with which they could interact was limited, thus increasing the possibility of interacting with the tool (a rake) that was presented on a table in front of them. It is, therefore, possible that the monkey succeeded not because of 'insight', but because this situation highly facilitated trial-and-error learning. Our data indicate that in the absence of any physical restraint and with a relatively easy sensorimotor task, the monkeys did not learn the use of the tool and could not accidentally obtain the food, a key factor in producing associative learning through operant conditioning.

After the shaping procedure, the monkeys became capable of using the stick. However, when presented with tools of different shapes (irrelevant feature for task accomplishment) and different length (relevant feature), both monkeys did not select the tool of adequate length for retrieving food in the very first trials. Instead, they first randomly chose a tool, grasped it and tried to use it for retrieving food. If the first attempt succeeded, then they continued to use that tool more frequently, otherwise they selected by chance another one. Despite the random choice observed in the first trial, overall the monkeys demonstrated a more frequent use of the functional tool. These results might, therefore, reflect a very rapid trial-and-error learning process.

In the condition in which all three tools were functional, monkeys (especially M1) tended to choose more frequently the stick which they had previously learned to use. It is possible that, in this condition, the familiarity with the stick based on previous sensorimotor experience facilitated its use, in line with previous data [14]. This issue could have been solved by presenting the monkey with three completely new tools, in order to verify its capacity for actively selecting the functional one regardless of its similarity to previously employed tools. Some observations made before the formal experiment on M1, after it learned to use the stick, showed that when the monkey was presented with only one unfamiliar tool, it demonstrated the capacity to generalize the use of the tool to the new objects despite their different and novel shapes. This supports the idea that monkeys possess a certain capacity to select objects as tools based on their general physical properties and regardless of their familiarity. Future experimental studies should investigate specifically the effect of familiarity on the capacity of tool selection.

From the motor point of view, the fact that the monkeys applied the same pattern of arm extension and tool insertion into the container when using the unfamiliar tools suggests that, despite the fact that they were not capable of identifying the correct physical properties which differentiate functional from non-functional tools, they could generalize the tool-use strategy previously learnt with one object to new objects, differing in shape, size and weight. When facing novel situations, they seem capable of rapidly learning these new associations.

Which factors might contribute to such a generalization process? Study 2 was aimed at investigating this issue in more depth.

(b) *Are generalization capacities of tool use based on a comprehension of the means-ends relationship?*

The two macaque monkeys employed in study 1 were presented with a new contextual situation, requiring them to use the stick with a completely new motor strategy.

At the beginning of the Baseline phase, both monkeys approached the container and attempted to extract the food by inserting their fingers or by biting it. Subsequently, they also made several attempts to retrieve food with the tool. Although M1 made fewer attempts than M2, it still used the stick and interacted with it in proximity to the container. Furthermore, both monkeys persisted in performing these manipulative behaviours for a long time. These observations induce us to interpret the monkeys manipulative behaviours as rudimentary attempts to use the tool, thus suggesting that they did know *what* to do, but not *how* to do it. Together, these findings also indicate that both monkeys 'conceived' the tool as a means to achieve the goal and that, once they had retrieved their neural motor representation for tool use, they tried to adjust it to a new context.

The difficulty observed in accomplishing the task with new motor strategies could be due to the previous sensorimotor practice and prolonged training which they experienced during study 1. This long-lasting training, in which the same tool-use behaviour was repeated in several sessions, could have resulted in reinforcing the link between the behavioural patterns used by the monkey and the type of target stimulus, thus favouring motor stereotypy. In support of this view, the results of study 2 showed that one of the two monkeys (M1), in the very first stages of the experiment, attempted to insert the tool into the container

by applying the same motor patterns (arm extension and wrist rotation/flexion) learnt during study 1.

During the Familiarization/Facilitation phase, after the tool was inserted by the experimenter into the container, both monkeys persistently attempted to insert the tool into the container in spite of their repeated failures. These repeated attempts could be based on a primitive and associative form of representation of means-ends relationship. Thus, after having extracted the tool with the food at the beginning of this phase, the monkeys could have transferred to the new contextual situation the link between the tool and the consequence of its use in the presence of a container with food. The final posture of the forelimb before the stick extraction movement could have facilitated the creation of a new motor pattern linking the forelimb sensorimotor representation occurring during the attempt, with the posture taken during grasping of the inserted tool. Moreover, the fact that this posture and tool extraction was followed by a reward, very likely increased the number of attempts and the probability to interact successfully with the container. In agreement with our interpretation, a study on free-ranging Japanese macaques showed that some individuals started to insert a stick into a pipe in order to retrieve the food trapped inside it after they were trained to pull the tool that was previously inserted by the experimenter [6]. Although, in that study, there was no description of the attempts and of the motor patterns used by the monkeys, factors similar to those described in the present study might have played a role.

The number of successful insertions rapidly increased within three sessions following the Familiarization/Facilitation phase. Although the monkeys had a limited time of exposure to the task, it is likely that the high success rate they obtained in the following practice phase depended, at least in part, on the time spent in manipulating the tool in proximity of the container in this phase.

Once the monkeys had learned how to perform the task, we directly assessed whether their behaviour was actually guided by a representation of the behavioural goal. In the Food/No Food test, we verified that monkeys rigidly applied the newly acquired motor strategy of tool use regardless of whether the food was present or not in the container. Probably, a further trial-and-error learning process enabled them to extinguish tool-use behaviour when the action did not lead to the reward. This lack of behavioural flexibility, as also evidenced in the first part of study 2, could be due to the high frequency of repetition of the previously learned behaviour, and can also be well framed within a neurophysiological perspective.

It has been demonstrated that neurons in areas of the monkey parietal and premotor cortices undergo changes as a consequence of tool use [16,32,33]. In one study, in which monkeys were trained to use different tools to grasp food morsels, it has been found that motor neurons of the ventral premotor cortex, normally active during hand grasping, also fired when the monkeys grasped the food with the tools, regardless of the exact movement sequence required for the purpose [16]. This finding clearly exemplifies that hand movements and associated goals are tightly linked in a unitary representation at the single neuron level, that is referred to as a 'motor act' [35]. As a consequence, the monkeys in our study very likely have motor representations in which goals and means are tightly linked to each other, and therefore cannot be processed independently. This could justify why the apparent monkeys' knowledge of the motor goal cannot be flexibly used to adapt their motor pattern to a novel context and to take into account the absence/presence of reward. In another study [32], authors investigated the properties of bimodal (somatosensory and visual) neurons of the parietal cortex in monkeys trained to use a rake to retrieve food out of reach. The results showed that the use of tools modified the size of the peripersonal receptive fields of the studied bimodal neurons and their modifications depended on whether the monkey actively used the tool. Altogether, these and other findings [17] suggest that the use of tools modifies the body schema and generates new motor representations, as if the tool becomes part of the body. Although the creation of new motor representations is indicative of brain plasticity, these representations tend to be rigidly used in strict relation to the tool-use behaviour in which they have been created.

(c) *Social learning processes of tool use: from behavioural data to possible neurophysiological mechanisms*

There is no clear consensus in the literature about the social learning abilities of macaques. While some studies showed that after a relatively long exposure to a demonstrator using a tool other individuals rarely acquire the new behaviour [8,10,11,36], several other studies have reported the social transmission and maintenance of novel behaviours in macaques [7,37], a phenomenon that is considered very important in creating new cultural achievements [38,39], especially where behaviours relative to food habits are concerned [40].

In the present study, we explored this issue in adult rhesus macaques with the aim of clarifying the social factors and cognitive processes possibly underlying the learning of tool use in this species. The main evidence provided by our study contrasts with the idea that macaques can learn tool use through a rapid observational learning process. Although during the Observation-Delayed tool interaction phase observers increased the frequency of tool manipulation, they never attempted to use the tool.

Beyond the possible cognitive limitations affecting the monkeys' capability to learn by observation from conspecifics, other factors may further contribute in explaining such failure. Among these factors, social tolerance is crucial since it allows individuals to observe others in close proximity and to directly participate in their activities. Rhesus macaques, instead, are known to be socially intolerant, especially during food processing [41]. Although they gather together while foraging, they do not tolerate the close proximity of other individuals, members of the same troop. Our data seem to support an impact of social intolerance on observational learning. In fact, when observers could use the tool simultaneously with the model,

they sometimes seemed to be inhibited by the model. This behavioural inhibition could reflect either a species-specific trait or the lack of social relations between the model and observer. We also cannot exclude the occurrence of behavioural extinction since, in the last few sessions of the Observation-Delayed tool interaction period, observers decreased their hand interactions with the tool.

Finally, another factor that could have contributed to the poor social learning abilities showed by the monkeys in the present study is the perspective from which they viewed the demonstrator using the tool, that is, a third-person perspective. To observe an action in a third-person rather than a first-person perspective implies additional cognitive operations such as mental rotation and transformation of the perceptual appearance of the observed action into a correspondent motor plan. Interestingly, it has been recently demonstrated that observing an action from first- or third-person perspective can activate different subpopulations of mirror neurons in the ventral premotor cortex [42]. Furthermore, other neurophysiological studies showed that when two monkeys sitting near each other are allowed to interact in a competitive situation, parietal neurons present complex combinatorial responses to 'of self' and 'other's' motion [43]. It is, therefore, possible that these neural mechanisms enable the monkeys to exploit the sight of other's action from several perspectives in order to better organize an appropriate response.

The behavioural and cognitive processes responsible for social learning in macaques have for long been at the centre of debate. Complex phenomena such as some form of imitation have been shown in macaques and other monkeys [44,45], and some authors argued that similar mechanisms may underlie tool-use learning in capuchin monkeys and chimpanzees [46]. Conversely, others have proposed that mechanisms different from imitation, such as stimulus or social enhancement, play a major role in promoting the individual discovery of how to use a tool [21,23].

In the current study, although monkeys did not learn by observation the use of the tool, they were clearly facilitated in interacting with it by observing the model's action. This effect is quite robust, but in the absence of any apparent reward deriving from the manipulative activity, the interaction with the tool decreases after a few sessions. Despite the social enhancement of manipulative behaviours with the tool during the Observation-Delayed tool interaction, naive observers never attempted to use it. From the observer's perspective, the model's activities with the tool might have enhanced the salience of the object, thus affecting its visual attention and interest towards it. However, considering the time lag between the observation and the execution phase, this explanation seems unlikely. Despite not being mutually exclusive with the stimulus-enhancement-hypothesis, it is also possible that the observation of grasping actions *per se* exerts facilitating effects on the observers' behaviour. For example, previous studies have demonstrated that in macaques and capuchin monkeys, the observation of and the listening to eating actions facilitate the performance of the same actions [47,48].

The presence of mirror neurons in the monkey motor cortex has prompted the idea that monkeys, as well as humans, understand the goal of an observed action by mapping its visual description onto the corresponding cortical motor representation [49]. According to this view, motor representation in the observer's brain enables him/her to directly understand the behavioural goal of the acting agent.

In the present experiment, it is very likely that observer monkeys could understand the goal of the model's action when it grasped the tool and when it licked it to eat the yogurt. All these actions were familiar to the observers and were part of their motor repertoire. Do they also have an understanding of the action made with the tool? Clearly, when monkeys are exposed to the observation of an action which they master because of a prolonged sensorimotor training, mirror neurons in their premotor cortex fire during observation of actions performed with the tool [50]. More interestingly, after prolonged visual exposure to an action performed by an experimenter using a tool, some premotor mirror neurons have been shown to respond specifically to the observation of tool actions [47]. In this latter study, however, when tested in their home cage, monkeys were not capable of using the same tool to which they were previously visually exposed. This is in line with the present findings showing an increased interaction of the observed monkey with the tool in the Observation-Delayed tool interaction phase. However, the lack of evidence of tool use indicates that monkeys do not have the ability to translate the visual description of the observed unfamiliar action into the motor programmes necessary for copying its behavioural goal.

6. CONCLUSIONS AND EVOLUTIONARY REMARKS

The literature on tool use shows that there is a discontinuity among different primate species, such as apes, capuchin monkeys and macaques. Macaques appear not to be proficient tool users, as testified by the very few reports on this topic in the wild [7]. Nevertheless, there is consistent evidence that, in this genus, the plasticity of the motor system is such to include tools as part of its 'vocabulary' of motor representations. In fact, not only can macaques be trained to use a variety of tools, but their use can be partially generalized to other objects and contexts. Interestingly, Iriki and co-workers [33,51] showed direct evidences of tool use-induced anatomical modifications in the temporal and parietal cortices, and the development of new cortico-cortical connections. Furthermore, this plasticity process appears to involve regions that are crucial for hand grasping [52]. In an evolutionary perspective, it is possible that cortical areas more susceptible to modifications as a result of tool use became more specialized for this function and separated from those just involved in sensorimotor transformation for hand grasping [53,54], supporting the idea that the use of tools required brain changes that determined the appearance of a new network.

An important capacity underlying tool use is that of combining single motor acts in action sequences.

The construction of complex action sequences implies neural structures capable of dealing with and integrating spatial and temporal features. For example, several studies reported the use of hammers and anvils in order to crack nuts in capuchin monkeys [55,56] and chimpanzees [57,58]. These activities may require complex behavioural patterns, such as the selection of an efficient hammer, its transportation to the location of the anvil, and the choice of the most appropriate motor pattern (trajectory, force, etc.). Thus, the brain of these primates is involved in several cognitive operations: individuation of a final goal, planning of the whole motor sequence necessary to reach this goal, a mental representation of the goal even in the actual absence of the sensory elements that drive the final part of the action sequence. This mental representation involves the capacity to travel in both space and time. Neurophysiological investigations have demonstrated the presence of circuits involved in motor planning and organization of sequential actions [59–61].

The issue of sequential organization of behaviour extends beyond the use of tools and embraces several other domains and, among them, speech. In fact, in speech, sequential organization is very important both in phonological articulation and for building a syntactic structure. Although the neural mechanisms underlying these processes can be the subject of investigation only in humans, many anatomical and functional data suggest that the neural substrates of sequential organization in non-human primates have provided the raw material for extending the properties of the cortical motor system to the domain of articulatory speech [62–65].

From the anatomical point of view, the rostral part of the macaque ventral premotor cortex has been considered, on the basis of anatomical location and cytoarchitectonic properties, as homologous to part of the human Broca's area. Functionally, neuroimaging studies in humans demonstrated that this latter area activates not only during speech production, but also during execution and observation of mouth and hand motor acts [63]. The same motor areas (ventral premotor cortex and posterior part of inferior frontal gyrus) involved in speech production and hand/mouth action organization seem to play an important role in tool use [66–68]. This brain regional overlap suggests that a basic organization of the motor system for hand and mouth actions has been exploited for the emergence of new functions that nonetheless rely, at least in part, on the same mechanisms.

REFERENCES

1 Ottoni, E. B. & Izar, P. 2008 Capuchin monkey tool use: overview and implications. *Evol. Anthropol.* **17**, 171–178. (doi:10.1002/evan.20185)

2 Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P. & Fragaszy, D. 2009 Selection of effective stone tools by wild bearded capuchin monkeys. *Curr. Biol.* **19**, 213–217. (doi:10.1016/j.cub.2008.11.064)

3 McGrew, W. C. 1992 Chimpanzee material culture: implications for human evolution. Cambridge, UK: Cambridge University Press.

4 Fragaszy, D., Izar, P., Visalberghi, E., Ottoni, E. B. & De Oliveira, M. G. 2004 Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *Am. J. Primatol.* **64**, 359–366. (doi:10.1002/ajp.20085)

5 Boesch, C., Head, J. & Robbins, M. 2009 Complex tool sets for honey extraction among chimpanzees in Loango National Park, Gabon. *J. Hum. Evol.* **56**, 560–569. (doi:10.1016/j.jhevol.2009.04.001)

6 Tokida, E., Tanaka, I., Takefushi, H. & Hagiwara, T. 1994 Tool-use in Japanese macaques: use of stones to obtain fruit from a pipe. *Anim. Behav.* **47**, 1023–1030. (doi:10.1006/anbe.1994.1140)

7 Gumert, M. D., Kluck, M. & Malaivijitnond, S. 2009 The physical characteristics and usage patterns of stone axe and pounding hammers used by long-tailed macaques in the Andaman Sea region of Thailand. *Am. J. Primatol.* **71**, 594–608. (doi:10.1002/ajp.20694)

8 Beck, B. B. 1976 Tool use in captive pigtailed macaques. *Primates* **17**, 301–310. (doi:10.1007/BF02382787)

9 Westergaard, G. C. 1988 Lion-tailed macaques (*Macaca silenus*) manufacture and use tools. *J. Comp. Psychol.* **102**, 152–159. (doi:10.1037/0735-7036.102.2.152)

10 Zuberbühler, K., Gygax, L., Harley, N. & Kummer, H. 1996 Stimulus enhancement and spread of spontaneous tool use in a colony of long-tailed macaques. *Primates* **37**, 1–12. (doi:10.1007/BF02382915)

11 Ducoing, A. M. & Thierry, B. 2005 Tool use in Tonkeana macaques (*Macaca tonkeana*). *Anim. Cogn.* **8**, 103–113. (doi:10.1007/s10071-004-0240-0)

12 Leca, J. B., Gunst, N. A., Watanabe, K. & Huffman, M. A. 2007 A new case of fish-eating in Japanese macaques: implications for social constraints on the diffusion of feeding innovation. *Am. J. Primatol.* **69**, 821–828. (doi:10.1002/ajp.20401)

13 Leca, J. B., Gunst, N. A. & Huffman, M. A. 2010 The first case of dental flossing by a Japanese macaque (*Macaca fuscata*): implications for the determinants of behavioral innovation and the constraints on social transmission. *Primates* **51**, 13–22. (doi:10.1007/s10329-009-0159-9)

14 Ishibashi, H., Hihara, S. & Iriki, A. 2000 Acquisition and development of monkey tool-use: behavioral and kinematic analyses. *Can. J. Physiol. Pharmacol.* **78**, 958–966. (doi:10.1139/y00-063)

15 Hihara, S., Obajashi, S., Tanaka, M. & Iriki, A. 2003 Rapid learning of sequential tool use by macaque monkey. *Physiol. Behav.* **78**, 427–434. (doi:10.1016/S0031-9384(02)01006-5)

16 Umiltà, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. 2008 When pliers become fingers in the monkey motor system. *Proc. Natl Acad. Sci. USA* **105**, 2209–2213. (doi:10.1073/pnas.0705985105)

17 Maravita, A. & Iriki, A. 2004 Tools for the body (schema). *Trends Cogn. Sci.* **8**, 79–86. (doi:10.1016/j.tics.2003.12.008)

18 Santos, L. R., Miller, C. & Hauser, M. D. 2003 Representing tools: how two non human primate species distinguish between the functionally relevant and irrelevant features of a tool. *Anim. Cogn.* **6**, 269–281. (doi:10.1007/s10071-003-0171-1)

19 Hauser, M. D. 1997 Artifactual kinds and functional design features: what a primate understands without

language. *Cognition* **64**, 285–308. (doi:10.1016/S0010-0277(97)00028-0)

20 Visalberghi, E. & Limongelli, L. 1994 Lack of comprehension of cause-effect relations in tool-using capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* **108**, 15–22. (doi:10.1037/0735-7036.108.1.15)

21 Fragaszy, D. & Visalberghi, E. 2004 Socially-biased learning in monkeys. *Anim. Learn. Behav.* **32**, 24–35. (doi:10.3758/BF03196004)

22 Nahallage, C. A. D. & Huffman, M. A. 2007 Acquisition and development of stone handling behavior in infant Japanese macaques. *Behaviour* **144**, 1193–1215. (doi:10.1163/156853907781890959)

23 Coussi-Korbel, S. & Fragaszy, D. M. 1995 On the relationship between social dynamics and social learning. *Anim. Behav.* **50**, 1441–1453. (doi:10.1016/0003-3472(95)80001-8)

24 Whiten, A. 2000 Primate culture and social learning. *Cogn. Sci.* **24**, 477–508. (doi:10.1207/s15516709cog2403_6)

25 van Schaik, C. P. 2003 Local traditions in orangutans and chimpanzees: social learning and social tolerance. In *The biology of traditions: models and evidence* (eds D. Fragaszy & S. Perry), pp. 159–186. Cambridge, UK: Cambridge University Press.

26 van Schaik, C. P., Fox, E. A. & Fechtman, L. 2003 Individual variation in the rate of use of tree-hole tools among wild orang-utans: implications for hominin evolution. *J. Hum. Evol.* **44**, 11–23. (doi:10.1016/S0047-2484(02)00164-1)

27 Perry, S. & Manson, J. H. 2003 Traditions in monkeys. *Evol. Anthropol.* **12**, 71–81. (doi:10.1002/evan.10105)

28 Lonsdorf, E. 2006 What is the role of mother in the acquisition of termite-fishing behaviours in the wild chimpanzees (*Pan troglodytes schweinfurthii*)? *Anim. Cogn.* **9**, 36–46. (doi:10.1007/s10071-005-0002-7)

29 Moscovice, L. R. & Snowdon, C. T. 2006 The role of social context and individual experience in novel task acquisition in cotton-top tamarins (*Saguinus oedipus*). *Anim. Behav.* **7**, 933–943. (doi:10.1016/j.anbehav.2005.09.007)

30 Range, F. & Huber, L. 2007 Attention of common marmosets: implications for social learning experiments. *Anim. Behav.* **73**, 1033–1041. (doi:10.1016/j.anbehav.2006.07.015)

31 Hirata, S., Watanabe, K. & Kawai, M. 2001 Sweet potato washing revisited. In *Primate origins of human cognition and behavior* (ed. T. Matsuzawa), pp. 487–508. Tokyo, Japan: Springer.

32 Iriki, A., Tanaka, M. & Iwamura, Y. 1996 Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* **14**, 2325–2330.

33 Hihara, S., Notoya, T., Tanaka, M., Ichinose, S., Ojima, H., Obayashi, S., Fujii, N. & Iriki, A. 2006 Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys. *Neuropsychologia* **44**, 2636–2646. (doi:10.1016/j.neuropsychologia.2005.11.020)

34 Van Hooff, J. A. R. A. M. 1967 The facial displays of Catarrhine monkeys and apes. In *Primate Ethology* (ed. D. Morris), pp. 7–68. London, UK: Weidenfeld & Nicholson.

35 Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. & Matelli, M. 1988 Functional organization of inferior area 6 in the macaque monkey. *Exp. Brain Res.* **71**, 491–507. (doi:10.1007/BF00248742)

36 Chauvin, C. & Berman, C. M. 2004 Intergenerational transmission of behavior. In *Macaque societies: a model for the study of social organization.* (eds B. Thierry, M. Singh & W. Kaumanns), pp. 209–234. Cambridge, UK: Cambridge University Press.

37 Leca, J., Gunst, N. & Huffman, M. 2010 Indirect social influence in the maintenance of the stone-handling tradition in Japanese macaques, *Macaca fuscata*. *Anim. Behav.* **79**, 117–126. (doi:10.1016/j.anbehav.2009.09.035)

38 Itani, J. & Nishimura, A. 1973 The study of intrahuman culture in Japan. A review. In *Precultural primate behavior* (ed. E. V. Menzel), pp. 26–55. Basel, Switzerland: Karger.

39 Huffman, M., Nahallage, C. A. D. & Leca, J. 2008 Cultured monkeys. Social learning cast in stones. *Curr. Dir. Psychol. Sci.* **17**, 410–414. (doi:10.1111/j.1467-8721.2008.00616.x)

40 Huffman, M. A. 1996 Acquisition of innovative cultural behaviours in nonhuman primates: a case study of SH, a socially transmitted behavior in Japanese macaques. In *Social learning in animals: roots of culture* (eds B. G. Galef Jr & C. Heyes Jr), pp. 267–289. San Diego, CA: Academic Press.

41 Thierry, B., Iwaniuk, A. N. & Pellis, S. M. 2000 The influence of phylogeny on the social behaviour of macaques (Primates: Cercopithecidae, genus Macaca). *Ethology* **106**, 713–728. (doi:10.1046/j.1439-0310.2000.00583.x)

42 Caggiano, V., Fogassi, L., Rizzolatti, G., Pomper, J. K., Thier, P., Giese, M. A. & Casile, A. 2011 View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Curr. Biol.* **21**, 144–148. (doi:10.1016/j.cub.2010.12.022)

43 Fujii, N., Hihara, S. & Iriki, A. 2007 Dynamic social adaptation of motion-related neurons in primate parietal cortex. *PLoS ONE* **2**, 1–8. (doi:10.1371/journal.pone.0000397)

44 Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggero, A. & Suomi, S. J. 2006 Neonatal imitation in rhesus macaques. *PLoS Biol.* **4**, 1501–1508. (doi:10.1371/journal.pbio.0040302)

45 Bugnyar, T. & Huber, L. 2000 True imitation in marmosets. *Anim. Behav.* **60**, 195–202. (doi:10.1006/anbe.2000.1457)

46 Whiten, A., McGuigan, N., Marshall-Pescini, S. & Hopper, L. M. 2009 Emulation, imitation over-imitation, the scope of culture for child and chimpanzee. *Phil. Trans. R. Soc. B* **364**, 2417–2428. (doi:10.1098/rstb.2009.0069)

47 Ferrari, P. F., Maiolini, C., Addessi, E., Fogassi, L. & Visalberghi, E. 2005 The observation and hearing of eating actions activates motor programs related to eating in macaque monkeys. *Behav. Brain Res.* **161**, 95–101. (doi:10.1016/j.bbr.2005.01.009)

48 Galloway, A. T., Addessi, E., Fragaszy, D. M. & Visalberghi, E. 2005 Social facilitation of eating familiar food in tufted capuchins (*Cebus apella*): does it involve behavioral coordination? *Int. J. Primatol.* **26**, 181–189. (doi:10.1007/s10764-005-0729-7)

49 Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)

50 Rochat, M., Caruana, F., Jezzini, A., Escola, L., Intskirveli, I., Grammont, F., Gallese, V., Rizzolatti, G. & Umiltà, M. A. 2010 Responses of mirror neurons in area F5 to hand and tool grasping observation. *Exp. Brain Res.* **204**, 605–616. (doi:10.1007/s00221-010-2329-9)

51 Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K., Lemon, R. N. & Iriki, A. 2009 Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proc. Natl Acad. Sci. USA* **106**, 18 379–18 384. (doi:10.1073/pnas.0909751106)

52 Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., Onoe, H. & Iriki, A. 2001 Functional brain mapping of monkey tool use. *Neuroimage* **14**, 853–861. (doi:10.1006/nimg.2001.0878)

53 Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G. & Orban, G. A. 2009 The representation of tool use in humans and monkeys: common and uniquely human features. *J. Neurosci.* **29**, 11 523–11 539.

54 Johnson-Frey, S. H. 2004 The neural bases of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78. (doi:10.1016/j.tics.2003.12.002)

55 Visalberghi, E., Spagnoletti, N., Ramos da Silva, E. D., Andrade, F. R., Ottoni, E., Izar, P. & Fragaszy, D. 2009 Distribution of potential suitable hammers and transport of hammer tools and nuts by wild capuchin monkeys. *Primates* **50**, 95–104. (doi:10.1007/s10329-008-0127-9)

56 de Resende, B. D., Ottoni, E. B. & Fragaszy, D. M. 2008 Ontogeny of manipulative behavior and nut-cracking in young tufted capuchin monkeys (*Cebus apella*): a perception-action perspective. *Dev. Sci.* **11**, 82–400. (doi:10.1111/j.1467-7687.2008.00731.x)

57 Inoue-Nakamura, N. & Matsuzawa, T. 1997 Development of stone tool use by wild chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **111**, 159–173. (doi:10.1037/0735-7036.111.2.159)

58 McGrew, W. C. 2010 In search of the last common ancestor: new findings on wild chimpanzees. *Phil. Trans. R. Soc. B* 27, 3267–3276. (doi:10.1098/rstb.2010.0067)

59 Tanji, J. & Hoshi, E. 2008 Role of the lateral prefrontal cortex in executive behavioral control. *Physiol. Rev.* **88**, 37–57. (doi:10.1152/physrev.00014.2007)

60 Rozzi, S., Ferrari, P. F., Bonini, L., Rizzolatti, G. & Fogassi, L. 2008 Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectony areas. *Eur. J. Neurosci.* **28**, 1569–1588. (doi:10.1111/j.1460-9568.2008.06395.x)

61 Fogassi, L., Ferrrai, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)

62 Rizzolatti, G. & Arbib, M. G. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194. (doi:10.1016/S0166-2236(98)01260-0)

63 Fogassi, L. & Ferrari, P. F. 2007 Mirror neurons and the evolution of embodied language. *Curr. Dir. Psychol. Sci.* **16**, 136–141. (doi:10.1111/j.1467-8721.2007.00491.x)

64 Gentilucci, M. & Corballis, M. C. 2006 From manual gesture to speech: a gradual transition. *Biobehav. Rev.* **30**, 949–960. (doi:10.1016/j.neubiorev.2006.02.004)

65 Pulvermüller, F. & Fadiga, L. 2010 Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* **11**, 351–360. (doi:10.1038/nrn2811)

66 Goldenberg, G., Hermsdörfer, J., Glindermann, R., Rorden, C. & Karnath, H. O. 2007 Pantomime of tool use depends on integrity of left inferior frontal cortex. *Cereb. Cortex* 7, 2769–2776. (doi:10.1093/cercor/bhm004)

67 Kroliczak, G. & Frey, S. H. 2009 A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. *Cereb. Cortex* **19**, 2396–2410. (doi:10.1093/cercor/bhn261)

68 Jacobs, S., Danielmeier, C. & Frey, S. H. 2010 Human anterior intraparietal and ventral premotor cortices support representations of grasping with the hand or a novel tool. *J. Cogn. Neurosci.* **22**, 2594–2608. (doi:10.1162/jocn.2009.21372)

Research

The neural and cognitive correlates of aimed throwing in chimpanzees: a magnetic resonance image and behavioural study on a unique form of social tool use

William D. Hopkins[1,2,*], **Jamie L. Russell**[2] **and Jennifer A. Schaeffer**[2]

[1]*Department of Psychology and Neuroscience, Agnes Scott College, 141 E. College Avenue, Decatur, GA 30030, USA*
[2]*Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, GA 30322, USA*

It has been hypothesized that neurological adaptations associated with evolutionary selection for throwing may have served as a precursor for the emergence of language and speech in early hominins. Although there are reports of individual differences in aimed throwing in wild and captive apes, to date there has not been a single study that has examined the potential neuroanatomical correlates of this very unique tool-use behaviour in non-human primates. In this study, we examined whether differences in the ratio of white (WM) to grey matter (GM) were evident in the homologue to Broca's area as well as the motor-hand area of the precentral gyrus (termed the KNOB) in chimpanzees that reliably throw compared with those that do not. We found that the proportion of WM in Broca's homologue and the KNOB was significantly higher in subjects that reliably throw compared with those that do not. We further found that asymmetries in WM within both brain regions were larger in the hemisphere contralateral to the chimpanzee's preferred throwing hand. We also found that chimpanzees that reliably throw show significantly better communication abilities than chimpanzees that do not. These results suggest that chimpanzees that have learned to throw have developed greater cortical connectivity between primary motor cortex and the Broca's area homologue. It is suggested that during hominin evolution, after the split between the lines leading to chimpanzees and humans, there was intense selection on increased motor skills associated with throwing and that this potentially formed the foundation for left hemisphere specialization associated with language and speech found in modern humans.

Keywords: throwing; Broca's area; chimpanzees

1. INTRODUCTION

Visitors to the zoo are sometimes treated to the sight of chimpanzees throwing objects (often faeces or wet chow) at each other or at them. What most zoo visitors do not appreciate is the rarity with which throwing occurs in non-human animals. Save for a few unsystematic and anecdotal reports of throwing in monkeys and great apes [1–9], there is little evidence that throwing occurs in other animals [10]. Thus, throwing appears to have come under positive selection pressure in hominins. From an evolutionary standpoint, some have suggested that throwing may have offered many advantages to early hominins such as the ability to kill larger prey without putting oneself at risk of being wounded or killed [11]. The ability to kill large game for the purposes of nutrition while simultaneously minimizing one's personal risk of injury or death (i.e. increased survival) would have been selectively advantageous [12]. Among non-human primates, throwing has most frequently been observed in wild and captive chimpanzees, though there are reports in other great ape species [1,5,6,8] and, to a lesser extent, in monkeys [13–15]. Throwing has been described as a form of tool use in chimpanzees at a number of long-term field sites in Africa, including Gombe, Mahale, Bossou and the Tai forest [16–19]. In the wild and in captivity, throwing has mostly been recorded in the context of both inter and intraspecies agonistic encounters, although some have described it as a means of initiating play or communication [20–22].

From a psychological and neurological standpoint, aimed throwing is very interesting for several reasons. First, it has been hypothesized that some instances of aimed throwing by chimpanzees reflect foresight or future planning on the part of the apes, an ability often described as uniquely human [23]. For instance, Osvath [7] eloquently describes a zoo-living male

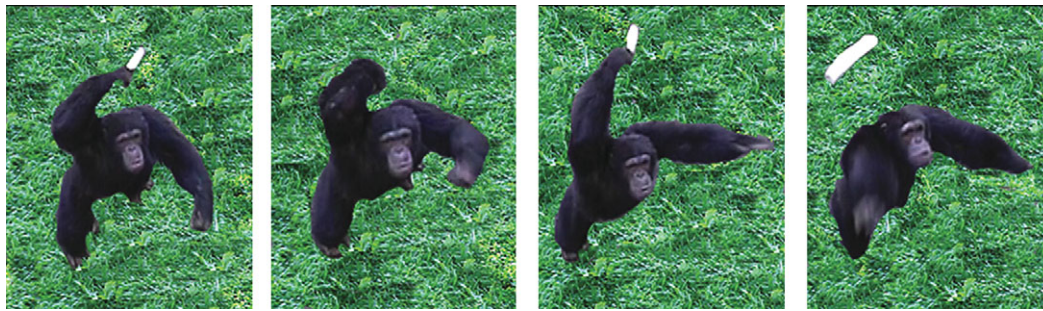* Author for correspondence (whopkin@emory.edu, whopkins@agnesscott.edu).

Figure 1. Four sequential frames demonstrating a chimpanzee throwing a polyvinylchloride (PVC) pipe towards a human in a tower above the subject. The chimpanzee stands bipedally, brings the PVC pipe back and then throws the object. Note that the force of the chimpanzees' throw causes him to leave the ground. This reflects the whole body function of throwing in some of the chimpanzees.

chimpanzee (named Santino) that hides rocks out of sight of the care staff, waiting to reveal and throw them at approaching visitors at the most opportune time. Evidence of planning comes from the observation that Santino searches for the rocks from a moat inside the enclosure prior to the arrival of the care staff and visitors, and caches the rocks out of sight, only to pull them out when the visitors arrive. We have made similar observations of this type of the so-called planning behaviour in the chimpanzees housed at the Yerkes National Primate Research Centre (YNPRC) and the University of Texas M. D. Anderson Cancer Centre [4]. Some of the chimpanzees will pile faeces or wet chow in their cage and wait for visitors to pass by before throwing this at them. We would further argue that aimed throwing in the YNPRC chimpanzees, though often agonistic in function and consequence, is not part of the apes' display behaviour. Indeed, most instances of aimed throwing that we have observed occur without any accompanying display behaviours such as pilo-erection, hooting or charging, further suggesting an element of planning on the part of the individual ape.

Throwing, as a form of social tool use, is also unique because it likely develops in captive chimpanzees (and possibly wild apes) by way of very different processes and reinforcement contingencies compared with other forms of tool use, notably those described for the purposes of food extraction. For instance, nut-cracking, termite-fishing and ant-dipping are by far the most common forms of tool use observed in wild chimpanzees and each of these is used for the purposes of obtaining otherwise unattainable food. Thus, in purely operant conditioning terms, the subjects learn to use these types of tools and maintain their use because they have been reinforced with food for successful use. Presumably, food is positively reinforcing and therefore increases the probability of subsequent occurrence of tool-use behaviour (though we would acknowledge that some forms of tool use may be maintained without explicit reinforcement).

In contrast, the rewards associated with throwing are quite different because they are not nutritive in form. Throwing in wild chimpanzees is seldom, if ever, observed for the purposes of obtaining food, but rather is almost always directed towards other chimpanzees or humans. In captivity, it is difficult to imagine that human caretakers would overtly reward a chimpanzee

with food immediately after they had just been soiled with faeces by the very same ape. In short, what appears to be the main reward for throwing is the simple ability to control or manipulate the behaviour of the targeted individual (ape or human). For example, in our laboratory, chimpanzees will patiently wait for strangers or visitors to approach and then will throw at them. They do not conceal their intentions and they will often stand bipedal and threaten to throw by cocking their arm with the projectile in their hand in preparation for throwing (figure 1). The passers-by can see this and will often try and negotiate with the chimpanzees to put down the projectile, or they will try to trick the ape by stopping, then dashing rapidly past the ape enclosure. This seems to be the reaction the apes hope to get from the humans and, in operant conditioning terms, is the only 'reward' the chimpanzees receive for throwing.

Neurologically, throwing is complex because it demands coordinated precision in timing the velocity and release window of a projectile in relation to the speed of movement and distance of the target (i.e. prey). Some have suggested that the increased selection for neural synchrony of rapid muscular sequencing routines associated with actions such as throwing are similar to the motor programming demands of language and speech, and therefore engage similar neural systems, notably Broca's area [24]. Moreover, because the left hemisphere is dominant for language, some have argued that the foundations for left hemisphere lateralization in language may have evolved from an initial preadaptation for right-handedness in throwing [11]. In Western cultures, a significant majority of individuals self-report preferring the right-hand for throwing [25], and studies in non-traditional societies have reported right-hand biases in throwing actions, such as in the use of spears [26]. Two previous studies have reported that captive chimpanzees show population-level right-handedness for throwing, which suggests left hemisphere dominance [4]. Hopkins *et al.* [4] have also found that posture influences handedness for throwing. Within the subsample of 89 chimpanzees that were observed to reliably throw by Hopkins *et al.* [4], 90 per cent of right-handed individuals preferred to throw overhand compared with underhand. A majority of the overhand throws were made when the chimpanzees were standing bipedally. In contrast, a significant majority of the left-handed individuals threw underhand

when they were in a quadrupedal posture. Thus, handedness was strongly linked to the posture and style of throwing adopted by the apes. Despite the long-standing theoretical interest in throwing in relation to brain evolution, few studies to date have examined the potential neural correlates of throwing in non-human primates, notably chimpanzees [27]. In a previous study, Cantalupo & Hopkins [27] found that chimpanzees that had learned to throw reliably had significantly larger cerebella than those that had not. Throwing offers a unique opportunity to consider cortical plasticity in chimpanzees because, as noted already, there are considerable individual differences in terms of its occurrence and lateralization.

One aim of this study was to examine whether variation in either grey matter (GM) or white matter (WM) within premotor and primary motor cortex was associated with the occurrence and lateralization of throwing in captive chimpanzees. Specifically, magnetic resonance image (MRI) scans were obtained in 76 chimpanzees and the proportions of WM to GM in the left and right hemispheres were computed for two cortical motor regions, including the motor-hand area of the precentral gyrus (termed the KNOB) and the inferior frontal gyrus (IFG). The KNOB was selected as a region of interest because it is the anatomical region of the precentral gyrus where the hand is represented, and studies in humans and chimpanzees have shown that asymmetries in this brain area are associated with hand preference [28–30]. Additionally, we measured the IFG because it is the homologue to Broca's area in the human brain [31,32] and previous studies in humans and chimpanzees have shown that variation in asymmetries in this region are associated with hand use for tool use [33–35]. If handedness for throwing is associated with asymmetries in these two motor regions, then we hypothesized that right-handed throwers would show leftward asymmetries while left-handed throwers would show rightward asymmetries. We further hypothesized that if learning to throw promotes the development of connections between cortical regions, then chimpanzees that reliably throw would show increased WM within the IFG and KNOB regions compared with individuals that have not learned to throw. We were particularly interested in WM because recent diffusion tensor imaging (DTI) studies have shown that training experiences can have a significant effect on cortical connectivity. For instance, Scholz *et al.* [36] measured WM connectivity in a sample of naive human participants and subsequently had them learn how to juggle. Post-training DTI imaging revealed increased cortical WM in several brain regions, but notably in regions underlying the intraparietal sulcus. More germane to this study, Quallo *et al.* [37] imaged three monkeys before and after training them on a tool-use task. Using voxel-based morphometry, significant increases in GM were found in several regions, including the intraparietal and superior temporal sulci. Furthermore, increased levels of WM were found bilaterally in the cerebellum—a brain structure that is critically involved in motor learning and coordination—after learning the tool-use task.

A second aim of this study was to test whether chimpanzees that have learned to throw are socially more sophisticated or intelligent than those that have not. As noted above, we and others have noted that chimpanzees that throw exhibit a considerable degree of planning in their actions and seem to know that they can use their throwing actions to manipulate or change the behaviours of other social beings, notably naive human observers. This suggests that these apes are more sensitive to how their actions influence the behaviours of others. Leavens *et al.* [38] argued that most ape gestural communication observed in captive individuals' functions in a similar manner. That is, when apes point to otherwise unattainable foods in the presence of humans, they are instrumentally controlling or manipulating the human to get the food for them, or in essence the human becomes the tool. For this reason, we hypothesized that apes that throw might be more socially or communicatively sophisticated than those that have not learned to throw. We tested this hypothesis by comparing the chimpanzees that reliably throw to non-throwers on a series of cognitive tasks that quantify physical and social cognition in apes. Specifically, Herrmann *et al.* [39] have described a series of tasks, referred to as the primate cognition test battery (PCTB), which allegedly measures the abilities of human children and non-human primates to use social and physical cues to solve different types of learning problems. Broadly, the tasks assess communication abilities, comprehension of causality, spatial cognition and memory, quantity discrimination and theory of mind. We have recently tested more than 90 chimpanzees on these tasks [40] and, in this study, we compared the performance of throwing and non-throwing apes to examine whether performance differences were evident in these groups.

2. METHODS

(a) Subjects

Magnetic resonance images were obtained from a sample of 78 chimpanzees, including 24 males and 54 females. The subjects ranged in age from 6 to 51 years (mean = 23.05, s.d. = 11.80). All the chimpanzees were members of a captive colony housed at YNPRC in Atlanta, GA, USA. Within the sample, there were 38 chimpanzees that reliably threw, and these individuals were matched on the basis of sex, age and scanning protocol with 38 chimpanzees that did not reliably throw. This was done to control for these potential confounding variables within the sample.

(b) Image collection and procedure

For the *in vivo* scanning, subjects were first immobilized by ketamine injection (10 mg kg^{-1}) and subsequently anaesthetized with propofol (40–60 mg (kg h^{-1})$^{-1}$) following standard procedures at the YNPRC. Subjects were then transported to the MRI facility. The subjects remained anaesthetized for the duration of the scans as well as the time needed to transport them between their home cage and the imaging facility (total time approximately 2 h). Subjects were placed in the scanner chamber in a supine position with their head fitted inside the human-head coil. Scan duration ranged between 40 and 60 min as a function of brain size.

A portion of the subjects were scanned using a 1.5 T scanner (Phillips, Model 51), while the remaining chimpanzees were scanned using a 3 T scanner (Siemens Trio, Siemens Medical Solutions USA, Inc., Malvern, Pennsylvania, USA) at the YNPRC.

For all chimpanzees scanned *in vivo* using the 1.5 T machine, T1-weighted images were collected in the transverse plane using a gradient echo protocol (pulse repetition = 19 ms, echo time = 8.5 ms, number of signals averaged = 8, and a 256 × 256 matrix). For the chimpanzees scanned using the 3 T scanner, T1-weighted images were collected using a three-dimensional gradient echo sequence (pulse repetition = 2300 ms, echo time = 4.4 ms, number of signals averaged = 3, matrix size = 320 × 320).

After completing MRI procedures, the subjects scanned *in vivo* were returned to the YNPRC and temporarily housed in a single cage for 6–12 h to allow the effects of the anaesthesia to wear off, after which they were returned to their social group. The archived MRI data were transferred to a PC running Analyze 7 (Mayo Clinic, Mayo Foundation, Rochester, MN, USA) software for post-image processing. Prior to data collection, two raters blind to the hemisphere and handedness of the chimpanzees independently measured the IFG and KNOB in 10 specimens. Inter-rater correlations between the two tracers were positive and significant for both regions [30,41]. Prior to measurement, the raw T1-weighted MRI scans were aligned in the axial, coronal and sagittal planes along the AC-PC line.

(i) Inferior frontal gyrus

The IFG was quantified separately for the left and right hemispheres in the axial (transverse) plane following procedures that have been described elsewhere [41] (figure 2a). The most dorsal point of the IFG slice was defined as the slice on which both the precentral inferior (PCI) and fronto-orbital (FO) sulci could be seen. Using a freehand tool, PCI was traced to the lateral portion of the brain and was followed until meeting FO. FO was then traced to its most medial point and the most medial points of FO and PCI were then connected with a straight line to create an estimate of the area of the gyrus between these two sulci for each slice. Both sulci had to be present to be considered a traceable slice. Successive 1 mm slices were traced using these landmarks until either FO or PCI were no longer visible. The areas within the traced regions were subsequently summed to derive volumes of the IFG for each hemisphere. These IFG object maps for the right and left hemispheres were saved for each individual subject and subsequent application to the segmented volumes.

(ii) Motor-hand area or KNOB

As with the IFG, the KNOB was quantified separately for the left and right hemispheres in the axial (transverse) plane (figure 2b) following procedures previously used in human and chimpanzee brain specimens [29,42]. The dorsal and ventral edges of the KNOB along the central sulcus (CS) served as markers
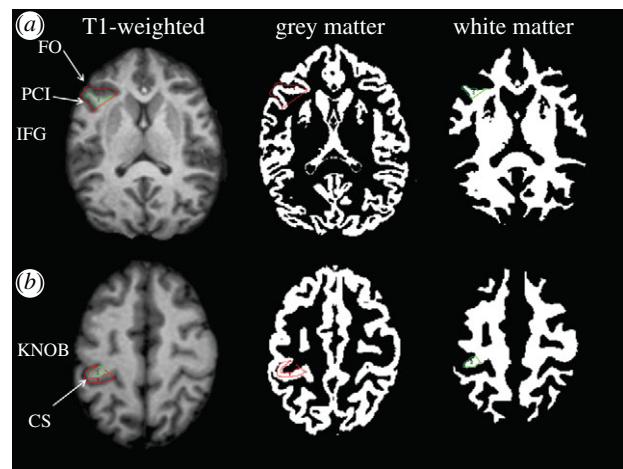
Figure 2. (*a*) T1-weighted axial view of a chimpanzee MRI scan followed by the segmented grey matter (GM) and white matter (WM) view. The landmarks used to quantify the IFG are indicated as well as the object maps as they were applied to the two segmented volumes. (*b*) Axial view of the KNOB region traced T1-weighted MRI scan with the object maps then applied to the GM and WM volumes. CS, central sulcus; FO, fronto-orbital sulcus; IFG, inferior frontal gyrus; PCI, precentral inferior sulcus.

for defining the boundaries of the area. The area of the entire KNOB was traced on each slice and hemisphere using a mouse-driven pointer (figure 2b). The areas within the traced regions were subsequently summed to derive volumes of the KNOB for each hemisphere. These KNOB tracings for the right and left hemispheres were saved for each individual subject.

(c) Image segmentation and region of interest measurements

The aligned T1-weighted MRI scans were skull-stripped and subsequently segmented into GM, WM and cerebral spinal fluid (CSF) tissue using FSL (Analysis Group, FMRIB, Oxford, UK) [43] (figure 2a,b). Because the segmented volumes were in the same stereotaxic space as the T1-weighted scans on which the object maps for the IFG and KNOB were drawn, the object maps were then applied to the segmented GM and WM, and the number of voxels that fell inside the object maps were calculated for each region, subject and hemisphere. We then divided the number of voxels by the total size of the object maps and multiplied by 100 to compute the percentage of GM or WM within each region and hemisphere. To simplify the analyses, we calculated WM-to-GM ratios for each hemisphere and region by dividing the WM percentage by the GM percentage (WM_GM ratio). Thus, the left and right hemisphere WM percentages for the KNOB and IFG were divided by the left and right GM percentages. We then computed average WM-to-GM ratios for the KNOB (WMGM_KNOB) and IFG (WMGM_IFG) by adding the values for the left and right hemispheres and dividing by two. We also computed asymmetry quotients (AQ) for the KNOB_AQ and IFG_AQ by using the formula $AQ = (R - L)/((R + L) \times 0.5)$, where R and L reflect the WM-to-GM ratios for the right and left hemispheres. Positive AQ values reflected right hemisphere

asymmetries, while negative values indicated left hemisphere biases.

(d) *Primate cognition test battery*
(i) *Procedures*
Subjects were tested on a modified version of the PCTB originally described by Herrmann *et al.* [39]. The PCTB attempts to assess subjects' abilities in various areas of physical and social cognition. For our study, some aspects of the original PCTB were eliminated owing to time and housing constraints. The previously published procedures were followed as closely as possible but some tasks were modified to better address the questions at hand, given the past experience and environmental constraints of our subjects. Each task is described briefly below with notes made when procedures were altered from those described by Herrmann *et al.* [39]. Subjects were generally tested in the order that the tasks are presented below and testing was completed over three to five testing sessions, depending on the motivation and attention of the subject. Most subjects were tested alone; however, some individuals are uncomfortable being separated from their group. These individuals were tested with one other conspecific with whom they were comfortable. All testing was done in the subject's home enclosure.

(e) *Physical cognition tasks*
Eight tasks were used in the 'Physical Cognition' portion of our test, including tasks exploring the apes' spatial memory and understanding of spatial relationships, ability to differentiate between quantities, understanding of causality in the visual and auditory domains and their understanding of tools. Our test differed from the original PCTB in several ways. We excluded the Addition task as well as certain components of the Tool Properties tasks.

(i) *Spatial memory* (*three trials*)
This test assessed subjects' ability to remember the locations of food rewards. In this task, the subject watched as food was hidden in two of three possible locations (opaque cups turned upside down) on a test table in front of them. The subject was then allowed to search the locations one by one. If the subject located one of two hidden food rewards, they were given the reward and allowed to search for the second hidden reward. Subjects that located both hidden food items without searching the unbaited location were scored as successful. Each subject received all three possible combinations of baited locations.

(ii) *Object permanence* (*nine trials*)
Here, we tested an individual's ability to follow a food reward after invisible displacement, given three different possible displacements. During single displacement trials, only one of three possible locations was manipulated and thus potentially baited. In the double displacement trials, two of three possible locations were manipulated, meaning that either location could potentially be baited. Double displacement trials were further divided by whether or not the baited locations

were adjacent to one another. In order to be considered successful, the subject must locate the hidden food item without searching in the location that was not manipulated.

(iii) *Rotation* (*nine trials*)
In the third task, subjects' ability to track a food reward as it is spatially rotated either 180° or 360° was examined. In this task, subjects watched as a food reward was hidden in one of the three locations (opaque cups turned upside down) lined up on a platform. The platform was then rotated on a horizontal plane, with the three locations being rotated as a unit. Three different manipulations were employed. In 180° middle trials, the middle location was baited and the platform was turned 180°. In 360° side and 180° side trials, either the left or right location was baited, and the platform was then rotated 360° or 180°, respectively. Subjects successfully completed a Rotation trial by tracking and identifying the correct location of the hidden reward.

(iv) *Transposition* (*nine trials*)
In this task, subjects watch as a food reward is hidden in one of three possible locations and then as the baited location is changed in one of the three ways. In one condition, the baited location is switched with one of the unbaited locations. In the second condition, the baited location is switched with one of the unbaited locations and then the two unbaited locations are switched. In the last condition, the baited location is switched with one of the unbaited locations and then with the other unbaited location. Subjects received three trials of each condition. To be considered successful on this task, the subject must track the reward and choose the baited location.

(v) *Relative numbers* (*13 trials*)
In the fifth task, subjects were tested for their ability to discriminate between different quantities by being presented with two plates containing different amounts of equally sized food pieces. Each subject received the same set of 13 different quantity pairings as those used in the original PCTB ($1:0$, $5:1$, $6:3$, $6:2$, $6:4$, $4:3$, $3:2$, $2:1$, $4:1$, $4:2$, $5:2$, $3:1$ and $5:3$). During each trial, the subject was allowed to choose only one plate and received whatever reward was on the chosen plate. A correct response was recorded when the subject chose the plate containing the larger quantity of food. We did not include the task by Herrmann *et al.* [39] referred to as Addition Numbers.

(vi) *Causality noise* (*six trials*)
In the sixth task, subjects' understanding of causal relationships based on sound was assessed. In this task, the experimenter placed a hard food reward (i.e. peanut) in one of the two metal containers such that the container with the food reward made a sound when shaken, while the unbaited container did not. In 'Full' trials, the metal container containing the food reward was lifted and shaken and then the unbaited container was lifted. In the 'Empty' trials, the empty container was lifted and shaken and then

the baited container was lifted. Subjects were then allowed to choose one of the two containers. A correct choice was recorded when the subject chose the baited container.

(vii) *Causality visual (six trials)*
In the seventh task, subjects were tested for their causal understanding of the physical world in the visual domain. Specifically, in one trial type, a food reward was placed underneath one of two boards lying flat on the testing table. The food caused the baited board to be tilted, while the unbaited board remained flat. In the second trial type, a food reward was placed underneath one of two pieces of cloth laying flat on the testing table. The reward created a visible bump in the baited cloth, while the unbaited cloth lay flat. Subjects received three trials with both the board and cloth. In both trial types, the subject had to choose the baited item to be considered successful.

(viii) *Tool properties (six trials)*
The last Physical Cognition task explored the apes' understanding of the physical properties of tools and how those relate to achieving a goal. In each task, the subject is presented with a choice between two similar tools. However, one tool can be used to obtain a food reward, while the other tool is ineffective. For the first task, subjects are presented with two identical pieces of paper. One piece of paper has a food reward sitting on top of the far end, while the second piece of paper has a food reward sitting beside it. The subject could pull either piece of paper into their cage, but only by pulling the paper with the food sitting on top of it would they be able to retrieve the food reward. In the second task, one tool was identical to the effective tool in the first task. The second tool consisted of two smaller pieces of paper with a small gap between them, visually emphasizing that they are disconnected. The food reward is placed on the out-of-reach piece of the two disconnected pieces of paper. The subject could pull in the reward using the effective tool, but pulling the piece of the disconnected paper is ineffective in obtaining the reward. Each subject received three trials for each tool property task. Note that we did not include three tool properties tasks from the original PCTB, 'Bridge', 'Broken Wool' and 'Tray Circle' [39].

(f) *Social cognition tasks*
The tasks designated as 'Social Cognition' in the PCTB that we used are fourfold. The first two are designed to test the apes' ability to understand and to produce communicative signals. The third set of tasks assesses their sensitivity to the attentional state of an experimenter and their ability to use appropriate communicative modalities based on this information. The last social cognition task is designed to assess rudimentary aspects of Theory of Mind by testing their ability to follow gaze. Owing to housing and time constraints, we excluded the Social Learning tasks done in the original PCTB and made some modifications to several of the other Social Cognition tasks noted as follows.

(i) *Comprehension (six trials)*
For the first task, we chose a slightly different strategy from that of Herrmann *et al.* [39] to assess the apes' abilities to comprehend communicative signals. The original task implemented the same table-and-cup set-up as used in many of the other physical cognition tasks. However, in our task, a target was placed on the left and right sides of the enclosure, while the subject was centred. The experimenter then used either gaze (three trials) or gaze combined with a manual point (three trials) to direct the subject to one of the two targets. The subject had to move to and touch the designated target to be considered successful on this task. Note that we did not include the 'Mark' condition conducted by Herrmann *et al.* [39].

(ii) *Production (four trials)*
For the second task, following the established PCTB methods, the apes' ability to produce communicative signals to indicate a hidden food item was tested in four trials [39]. In this task, the ape watched as an experimenter baited a location on either the far left or far right side of the enclosure. A second experimenter then approached the cage, centred the subject and waited for the subject to indicate which location contained the hidden food. The subject was given 60 s to indicate the correct location using an overt communicative signal, such as a manual gesture towards the hidden food.

(iii) *Attentional state (eight trials)*
In the third task, we followed the methods outlined by Herrmann *et al.* [39] but added an additional test. First, an experimenter placed a piece of food on the ground outside of the subject's enclosure. Then a second experimenter approached the cage and altered their attentional state in one of four ways. In the first trial, the experimenter's face and body were directed towards the food item and the subject. In the second trial, the experimenter's body faced the subject, but her face was turned away. In the third trial, the experimenter stood with her body facing away from the enclosure, but then turned her head to look at the subject. In the last trial, the experimenter's body and face were oriented away from the subject. In order to be successful, the subject had to use a communicative signal in the modality appropriate to the experimenter's attentional state. For example, if the experimenter was looking at the subject, he/she could use a visual signal, such as a manual gesture to indicate the food. However, if the experimenter was facing away from the subject, the subject had to first use an auditory or tactile signal, such as a cage bang or a spit to get the attention of the experimenter and then once the experimenter was looking at him/her, use a visual signal to indicate the food. To further explore this topic, we added an additional set of four trials using the same basic conditions. However, the trials were conducted in a more familiar setting with the experimenter sitting at the testing table, placing a piece of food on the table and then carrying out the four variations of attentional state. The same requirements regarding modality-specific communication

were required for the subject to be considered successful on any given trial.

(iv) *Gaze following* (*three trials*)

For the last social cognition test, we examined each ape's ability to follow gaze in three trials. In this task, an experimenter sat on a stool approximately 1 m from the subject's enclosure. The experimenter captured the subject's attention and centred him/her by offering a piece of food. The experimenter then shifted her head and eyes to gaze at a point directly above her head for a period of 10 s. In order to be successful, the subject had to follow the gaze of the experimenter by looking upward. Note that we did not test our subjects on two of the gaze following tasks, 'Back' and 'Eyes', used in the original PCTB [39].

(v) *Data analysis*

The mean proportion of correct trials was calculated for each of the 12 tasks. From these data, we computed average performance scores for the five basic cognitive dimensions originally described by Hermann *et al.*, which include COMMUNICATION (Comprehension, Production and Attentional State), CAUSALITY (Visual Causality, Tool Properties and Noise), SPACE (Spatial Memory, Object Permanence, Rotation and Transposition), QUANTITY (Relative Numbers) and THEORY OF MIND (Gaze Alternation).

3. RESULTS

(a) *Neuroanatomical correlates*

In our initial analyses, we compared the proportion of WM to GM within the IFG and KNOB between males and females. These analyses were compared on both the WMGM_KNOB and WMGM_IFG and KNOB_AQ and IFG_AQ values. For both analyses, no significant main effects or interactions were found. Thus, male and female chimpanzees did not differ in the proportion of WM to GM nor in lateralization for the IFG and KNOB.

We next considered the influence of throwing on the proportion of WM to GM within the KNOB and IFG. For this analysis, we performed a mixed model analysis of variance (ANOVA) with region (WMGM_IFG, WMGM_KNOB) serving as the within-subject factor, while throwing classification (THROW+, THROW−) was the between-group factor. Significant main effects for region ($F_{1,74} = 41.02$, $p < 0.001$) and throwing classification were found ($F_{1,74} = 5.6820$, $p < 0.03$). The proportions of WM found within the region object maps were significantly higher for the KNOB compared with those for IFG. Moreover, for both regions, the proportion of WM was significantly higher in the THROW+ than in the THROW− chimpanzees. The mean WGGM_KNOB and WMGM_IFG in THROW+ and THROW− chimpanzees are shown in figure 3.

(b) *Neuroanatomical correlates of laterality in throwing*

We next considered whether interhemispheric differences in the ratio of WM to GM were found for left- and right-handed chimpanzees on the basis of
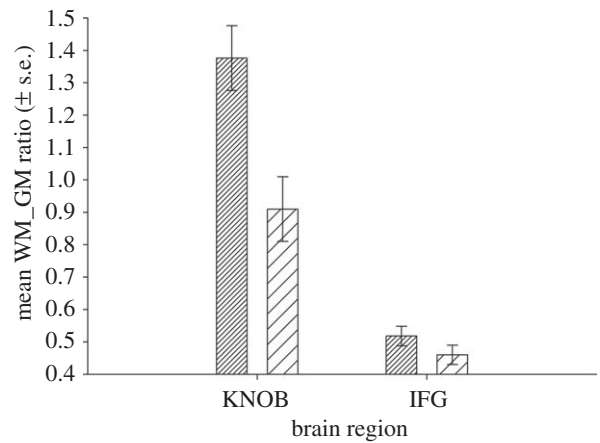
Figure 3. Mean white-to-grey matter ratio (Mean WM_GM Ratio) ($\pm$ s.e.) for chimpanzees that reliably throw (THROW+; narrow-striped bars) and those that do not (THROW−; wide-striped bars).
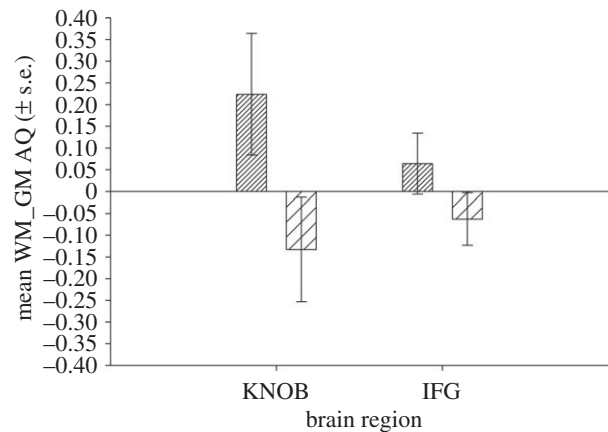
Figure 4. Mean WM_GM AQ scores ($\pm$ s.e.) for the IFG and KNOB in right- and left-handed throwing chimpanzees. Narrow-striped bars, left; wide-striped bars, right.

their preferred hand for throwing. For this analysis, we performed a mixed model ANOVA with the KNOB_AQ and IFG_AQ values serving as the within-subject factor, while throwing preference (LEFT, RIGHT) was the between-group factor. A significant main effect for throwing hand was found ($F_{1,36} = 5.462$, $p < 0.03$; figure 4). Right-handed throwers rather than left-handed throwers showed significantly greater leftward asymmetries in the WM-to-GM ratio for the IFG and KNOB.

(c) *Cognitive correlates of throwing*

As noted earlier, our laboratory has administered the PCTB test to 91 chimpanzees, and within this sample, there were 39 apes that threw consistently (THROW+) and 52 that did not (THROW−). In this next analysis, we compared THROW+ and THROW− individuals in their performance on the PCTB task. To test whether differences in cognitive abilities were evident between the two groups, we conducted a multiple analysis of variance (MANOVA) with the mean PCTB scores serving as dependent measures, while sex and throwing group served as the between-group factors. The MANOVA revealed a significant
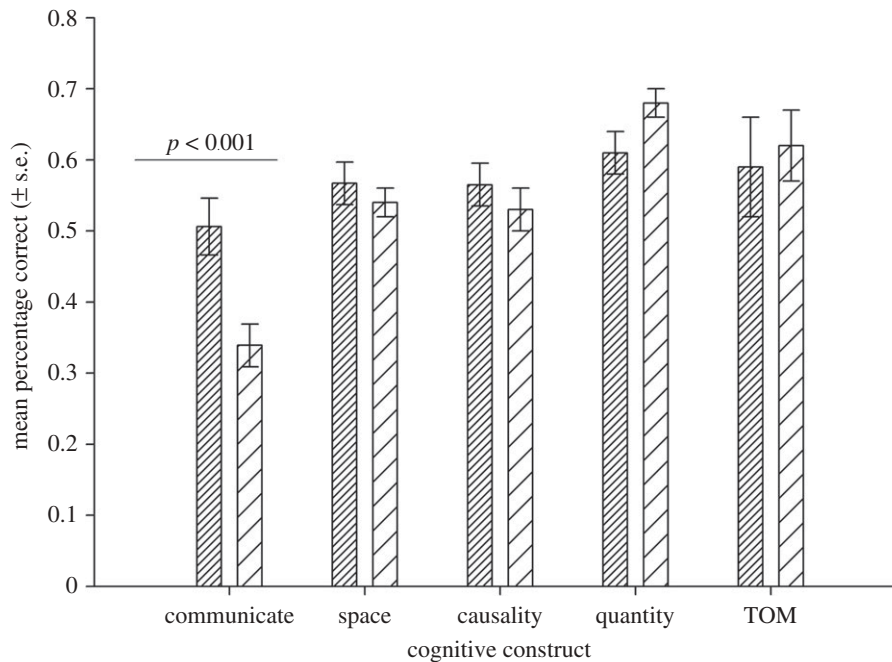
Figure 5. Mean percentage correct ($\pm$ s.e.) for the five cognitive abilities tested with the PCTB tasks in the chimpanzees (see text for description of the tasks). Narrow-striped bars, THROW+; wide-striped bars, THROW−.

main effect for throwing group ($F_{5,86} = 3.642$, $p < 0.001$). Subsequent univariate F-tests indicated that the only significant difference between the throwing groups was for COMMUNICATION ($F_{1,87} = 11.388$, $p < 0.001$) with THROW+ apes performing significantly better than the THROW− individuals (figure 5). No other significant differences were found.

4. DISCUSSION

In terms of the neuroanatomical correlates of throwing, the results reported here indicate that the presence of throwing skills in chimpanzees is correlated with increased cortical connectivity in Broca's area as well as in the motor-hand area of the precentral gyrus. The principal result in support of this conclusion is the observation that THROW+ apes had significantly higher proportions of WM to GM within the IFG and KNOB regions than THROW− individuals. Because increased WM indicates more myelinated interneurons that connect different cortical regions, this result would suggest that learning to throw may alter the connectivity between premotor and primary motor cortex in the chimpanzee. The results from comparing right- and left-handed throwers also support this conclusion, because differences in WM are hemisphere-specific and contralateral to the preferred hand that the chimpanzees use for throwing. The association between hand preferences for throwing and asymmetries within the IFG and KNOB are also consistent with previous studies showing that handedness for other forms of tool use, such as simulated termite-fishing and nut-cracking, is linked to lateralization in the cortical language area homologues of chimpanzees [33,44]. It should also be noted that asymmetries in hand use for manual gestures are associated with asymmetries within the IFG [45] Thus, the results reported here are consistent with the evolutionary hypothesis that

throwing may have served as a preadaptation for the neural adaptation of motor programmes necessary for complex motor actions, including language and speech [11].

It should also be noted that hand preferences for throwing are linked to variation in handedness for other measures of tool use and communication, but not non-tool-use measures of manual action. For instance, we have a large sample of hand preference data in chimpanzees for actions such as simple reaching, coordinated bimanual actions, simulated termite-fishing tool use and manual gestures [46–49]. Interestingly, right- ($n = 44$) and left-handed ($n = 37$) throwers differ significantly in their handedness for simulated termite fishing $t_{79} = -2.50$, $p < 0.02$ and manual gestures $t_{79} = -2.59$, $p < 0.02$, but not for simple reaching $t_{79} = -1.20$, n.s. or for coordinated bimanual actions $t_{79} = -0.62$, n.s. (figure 6). Thus, as with the neuroanatomical data, there is an explicit link between throwing, simulated termite fishing and gestural communication in terms of lateralization of function.

We also found that chimpanzees that have learned to throw are better at communication tasks than chimpanzees that have not. Interestingly, these two cohorts of chimpanzees do not differ on cognitive tasks that assess dimensions of physical cognition. These results suggest an explicit association between the cognitive foundations for throwing and the ability to engage in successful intraspecies communication, at least as assessed by the PCTB. Leavens *et al.* [38] have argued that, in captivity, chimpanzees learn to gesture to humans for foods that are otherwise unavailable to them by solving the referential problem space. That is, the chimpanzees want the food, but the food cannot be reached owing to physical barriers preventing the apes from attaining the foods. What the chimpanzees have learned to do with their gestures and other signals, such as attention-getting sounds,
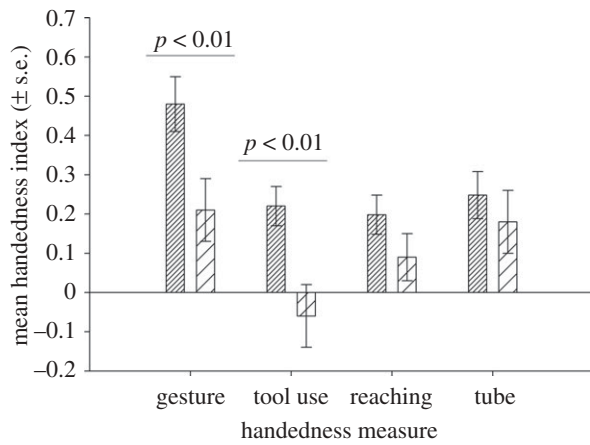
Figure 6. Mean handedness index (HI) scores ($\pm$ s.e.) for handedness measured in chimpanzees that prefer to throw with the right hand compared with those with the left. HI scores are computed by using the formula $AQ = (R - L)/R + L$), where R and L reflect the number of right- and left-hand responses. Positive HI values reflect right-hand preferences and negative values reflect left-hand preferences, respectively. Narrow-striped bars, THROW+; wide-striped bars, THROW−.

is to manipulate humans to obtain the food for them. Thus, in essence, the human becomes a tool for them. Indeed, we would argue that if a physical tool was available in these contexts, such as a long stick, the chimpanzees would use them to get the food rather than wait for a human caretaker to come by and retrieve the food for them.

Cognitively, we believe that the development and acquisition of throwing skills by chimpanzees operates in a manner similar to the emergence of manual gestural communication. As noted previously, the motivation for throwing in chimpanzees is largely to alter the behaviour of other individuals (be it human visitors or conspecifics). For this reason, the apes that have learned to throw have acquired an ability to understand how their behaviour affects the behaviours of others. If the same individuals apply these basic skills within the context of (i) understanding gestures, (ii) producing gestures, and (iii) using attention-getting behaviours when a human experimenter is inattentive to them, then it would follow that the THROW+ individuals should outperform the THROW− apes on the PCTB test. Moreover, these same skills may not offer any advantage in tasks that are not communicative in function.

We have focused on aimed throwing by chimpanzees in this paper, but some discussion of the prevalence of throwing in other species seems warranted within the context of the underlying cognitive processes that appear to accompany this ability in apes. In particular, capuchin monkeys have been reported to engage in aimed throwing [14] and they have well-documented tool-using abilities both in the wild and in captivity [50–56]. On the basis of our results in chimpanzees, this leads to the suggestion that capuchin monkeys might also engage in some forms of gestural communication during intraspecies interactions; however, this has not been frequently reported in the literature [57]. Thus, the

interconnection in social cognitive abilities between throwing and gestural communication may not be well developed in the capuchin monkey. We would also point out that, though the capuchin monkeys in the studies by Westergaard & Suomi [15] did learn to throw, their behaviour was explicitly shaped by the investigators by initially having the subjects throw an object into a bucket containing peanut butter (a preferred food). The peanut-butter-covered object was then handed to the subject and thereby the monkeys were reinforced with food for throwing. Thus, the acquisition of throwing by capuchin monkeys (at least in this study) appears to be mediated by a different type of reward when compared with the chimpanzees we have described in this paper.

There are limitations to the present study that warrant some discussion. First, it could be argued that the associations we found between throwing and WM reflect inherent differences in cortical organization rather than developing as a consequence of learning to throw. In other words, it may be the case that chimpanzees that learn to throw have inherently higher levels of WM within the KNOB and IFG regions and that this enabled them to learn to throw (rather than the increased WM emerging as a consequence of their experience). Similarly, though we found an association between throwing ability and communication skills on the PCTB tasks, it is not clear whether increased communication abilities are a consequence of learning to throw or vice versa. Neither of these alternative explanations can be ruled out based on the findings reported here.

One approach to address this issue would be to identify THROW− apes that differ with respect to WM volumes or asymmetries and then train them to throw. If inherent differences in WM explain the results reported here, then the prediction would be that individuals with larger WM volumes would acquire throwing abilities much faster than individuals with smaller volumes. In terms of communication, a similar approach could be used by performing a pre−post-test of communication abilities in two groups of apes, where one group is taught to throw and the other group is not. If learning to throw enhances communication abilities, then apes taught to throw should perform significantly better on the communication tasks in the post-test compared with apes that are not taught to throw.

Second, we examined the proportion of GM and WM within gyri comprising the regions of interest in this study (IFG and KNOB). This is a relatively crude measure of WM organization. Ideally, studies comparing THROW+ and THROW− apes using more sensitive measures of WM connectivity, such as DTI, would be more informative of the possible changes in cortical connectivity between the IFG and KNOB in these cohorts of apes [58,59].

In summary, we believe that this is the first evidence linking throwing with aspects of cortical organization and asymmetries in non-human primates as well as with differences in communicative abilities. The brain areas distinguishing right- and left-handed throwing chimpanzees show considerable overlap with cortical regions involved in language processing by humans.

Our findings that (i) chimpanzees are predominantly right-handed for throwing [4], and (ii) relative WM volumes within two frontal lobe regions distinguish right- from left-handed throwers suggest that the left hemisphere was specialized for the planning of complex motor actions prior to the split between the lines leading to humans and chimpanzees. Increasing selection for aimed throwing in a context of hunting or predator defence in hominins may have refined the neural architecture of the left hemisphere so as to eventually support other complex motor sequencing actions, including language and speech.

REFERENCES

1 Colell, M., Segarra, M. D. & Sabater-Pi, J. 1995 Hand preferences in chimpanzees (*Pan troglodytes*), bonobos (*Pan paniscus*) and orangutans (*Pongo pygmaeus*) in food-reaching and other daily activities. *Int. J. Primatol.* **16**, 413–434. (doi:10.1007/BF02735795)

2 Hamilton, W. J. I., Buskirk, R. E. & Buskirk, W. H. 1975 Defensive stoning by baboons. *Nature* **256**, 488–489. (doi:10.1038/256488a0)

3 Hopkins, W. D., Bard, K. A., Jones, A. & Bales, S. L. 1993 Chimpanzee hand preference in throwing and infant cradling: implications for the origin of human handedness. *Curr. Anthropol.* **34**, 786–790. (doi:10.1086/204224)

4 Hopkins, W. D., Russell, J., Cantalupo, C., Freeman, H. & Schapiro, S. 2005 Factors influencing the prevalence and handedness for throwing in captive chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **119**, 363–370. (doi:10.1037/0735-7036.119.4.363)

5 Jordan, C. 1982 Object manipulation and tool-use in captive pygmy chimpanzees (*Pan paniscus*). *J. Hum. Evol.* **11**, 35–39. (doi:10.1016/S0047-2484(82)80029-8)

6 Nakamichi, M. 1999 Spontaneous use of sticks as tools by captive gorillas (*Gorilla gorilla gorilla*). *Primates* **40**, 487–498. (doi:10.1007/BF02557584)

7 Osvath, M. 2009 Spontaneous planning for future stone throwing by a male chimpanzee. *Curr. Biol.* **19**, R190–R191. (doi:10.1016/j.cub.2009.01.010)

8 Parker, S., Kerr, M., Markowitz, H. & Gould, J. 1999 A survey of tool use in zoo gorillas. In *The mentalities of gorillas and orangutans: comparative perspectives* (eds S. T. Parker & R. W. Mitchell). Cambridge, UK: Cambridge University Press.

9 Takeshita, H. & Van Hooff, J. A. R. A. M. 1996 Tool use by chimpanzees (*Pan troglodytes*) of the Arnhem Zoo community. *Jpn. Psychol. Res.* **38**, 163–173. (doi:10.1111/j.1468-5884.1996.tb00021.x)

10 Wickler, W. & Seibt, U. 1997 Aimed object-throwing by a wild African elephant in an interspecific encounter. *Ethology* **103**, 365–368. (doi:10.1111/j.1439-0310.1997.tb00152.x)

11 Calvin, W. H. 1983 *The throwing Madonna: essays on the brain.* New York, NY: MacGraw-Hill.

12 Dunsworth, H. M., Challis, J. H. & Walker, A. 2003 Throwing and bipedalism: a new look at an old idea. In *Upright walking* (eds J. L. Franzen, M. Dohler & S. Moya-Sola). Frankfurt, Germany: Senckenberg Institute.

13 Leca, J. B., Nahallage, C. A. D., Gunst, N. & Huffman, M. A. 2008 Stone-throwing by Japanese macaques: form and functional aspects of a group-specific behavioral tradition. *J. Hum. Evol.* **50**, 989–998. (doi:10.1016/j.jhevol.2008.06.004)

14 Westergaard, G. C., Liv, C., Haynie, M. K. & Suomi, S. J. 2000 A comparative study of aimed throwing by monkeys and humans. *Neuropsychologia* **38**, 1511–1517. (doi:10.1016/S0028-3932(00)00056-7)

15 Westergaard, G. C. & Suomi, S. J. 1994 Aimed throwing of stones by tufted capuchin monkeys (*Cebus apella*). *Hum. Evol.* **9**, 323–329. (doi:10.1007/BF02435518)

16 Boesch, C. & Boesch, H. 1990 Tool use and tool making in wild chimpanzees. *Folia Primatol.* **54**, 86–99. (doi:10.1159/000156428)

17 Goodall, J. 1986 *The chimpanzees of Gombe: patterns of behavior.* Cambridge, MA: Harvard University Press.

18 Nishida, T., Matsusaka, T. & McGrew, W. C. 2009 Emergence, propagation or disappearance of novel behavioral patterns in the habituated chimpanzees of Mahale: a review. *Primates* **50**, 23–36. (doi:10.1007/s10329-008-0109-y)

19 Whiten, A., Horner, V. & de Waal, F. B. 2005 Conformity to cultural norms of tool use in chimpanzees. *Nature* **437**, 737–740. (doi:10.1038/nature04047)

20 Liebal, K., Muller, C. & Pika, S. (eds) 2005 *Gestural communication in nonhuman and human primates.* Amsterdam, The Netherlands: John Benjamins.

21 Pika, S., Liebal, K. & Tomasello, M. 2003 Gestural communication in young gorillas (*Gorilla gorilla*): gestural repertoire, learning, and use. *Am. J. Primatol.* **60**, 95–111. (doi:10.1002/ajp.10097)

22 Pika, S., Liebal, K. & Tomasello, M. 2005 The gestural repertoire of bonobos (*Pan paniscus*): flexibility and use. *Am. J. Primatol.* **65**, 39–61. (doi:10.1002/ajp.20096)

23 Suddendorf, T. & Corballis, M. 2007 The evolution of foresight: what is mental time travel and is it unique to humans? *Behav. Brain Sci.* **30**, 299–351.

24 Kimura, D. 1993 *Neuromotor mechanisms in human communication.* Oxford, UK: Oxford University Press.

25 Bryden, M. P. 1977 Measuring handedness with questionnaires. *Neuropsychologia* **15**, 617–624. (doi:10.1016/0028-3932(77)90067-7)

26 Raymond, M. & Pontier, D. 2004 Is there geographical variation in human handedness? *Laterality* **9**, 35–51.

27 Cantalupo, C. & Hopkins, W. D. 2009 The cerebellum and its contribution to complex tasks in higher primates: a comparative perspective. *Cortex* **46**, 821–830. (doi:10.1016/j.cortex.2009.10.004)

28 Hammond, G. 2002 Correlates of human handedness in primary motor cortex: a review and hypothesis. *Neurosci. Biobehav. Rev.* **26**, 285–292. (doi:10.1016/S0149-7634(02)00003-9)

29 Hopkins, W. D. & Cantalupo, C. 2004 Handedness in chimpanzees is associated with asymmetries in the primary motor but not with homologous language areas. *Behav. Neurosci.* **118**, 1176–1183. (doi:10.1037/0735-7044.118.6.1176)

30 Hopkins, W. D., Taglialatela, J. P., Russell, J. L., Nir, T. M. & Schaeffer, J. 2010 Cortical representation of lateralized grasping in chimpanzees (*Pan troglodytes*): a combined MRI and PET study. *PLoS ONE* **5**, e13383. (doi:10.1371/journal.pone.0013383)

31 Keller, S. S., Roberts, N. & Hopkins, W. D. 2009 A comparative magnetic resonance imaging study of the anatomy, variability and asymmetry of Broca's area in the human and chimpanzee brain. *J. Neurosci.* **29**, 14 607–14 616. (doi:10.1523/JNEUROSCI.2892-09.2009)

32 Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A., Stimpson, C. D., Erwin, J. M., Hof, P. R. & Sherwood, C. C. 2010 Broca's area homologue in chimpanzees (*Pan troglodytes*): probabilistic mapping, asymmetry and comparison to humans. *Cereb. Cortex* **20**, 730–742. (doi:10.1093/cercor/bhp138)

33 Hopkins, W. D., Russell, J. L. & Cantalupo, C. 2007 Neuroanatomical correlates of handedness for tool use in chimpanzees (*Pan troglodytes*): implication for theories on the evolution of language. *Psychol. Sci.* **18**, 971–977. (doi:10.1111/j.1467-9280.2007.02011.x)

34 Johnson-Frey, S. H. 2004 The neural basis of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78. (doi:10.1016/j.tics.2003.12.002)

35 Lewis, J. W. 2006 Cortical networks related to human use of tools. *Neuroscientist* **12**, 211–231. (doi:10.1177/1073858406288327)

36 Scholz, J., Klein, M. C., Behrens, T. E. J. & Johansen-Berg, H. 2009 Training induces changes in white-matter architecture. *Nat. Neurosci.* **12**, 1370–1371. (doi:10.1038/nn.2412)

37 Quallo, M. M., Price, C. J., Ueno, K., Asamizuya, T., Cheng, K., Lemon, R. N. & Iriki, A. 2009 Gray and white matter changes associated with tool-use learning in macaque monkeys. *Proc. Natl Acad. Sci. USA* **106**, 18 379–18 384. (doi:10.1073/pnas.0909751106)

38 Leavens, D. A., Hopkins, W. D. & Bard, K. A. 2005 Understanding the point of chimpanzee pointing: epigenesis and ecological validity. *Curr. Dir. Psychol. Sci.* **14**, 185–189. (doi:10.1111/j.0963-7214.2005.00361.x)

39 Herrmann, E., Call, J., Hernandez-Lloreda, M. V., Hare, B. & Tomasello, M. 2007 Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science* **317**, 1360–1366. (doi:10.1126/science.1146282)

40 Russell, J. L., Lyn, H., Schaeffer, J. A. & Hopkins, W. D. In press. The role of socio-communicative rearing environments on the development of social and physical cognition in apes. *Dev. Sci.*

41 Hopkins, W. D., Taglialatela, J. P., Meguerditchian, A., Nir, T., Schenker, N. M. & Sherwood, C. C. 2008 Gray matter asymmetries in chimpanzees as revealed by voxel-based morphometry. *NeuroImage* **42**, 491–497. (doi:10.1016/j.neuroimage.2008.05.014)

42 Yousry, T. A., Schmid, U. D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A. & Winkler, P. 1997 Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* **120**, 141–157. (doi:10.1093/brain/120.1.141)

43 Zhang, Y., Brady, M. & Smith, S. M. 2001 Segmentation of the brain MR images through hidden Markov random filed model and expectation-maximization algorithm. *IEEE Trans. Med. Imaging* **20**, 45–57. (doi:10.1109/42.906424)

44 Cantalupo, C., Freeman, H. D., Rodes, W. & Hopkins, W. D. 2008 Handedness for tool use correlates with cerebellar asymmetries in chimpanzees (*Pan troglodytes*). *Behav. Neurosci.* **122**, 191–198. (doi:10.1037/0735-7044.122.1.191)

45 Taglialatela, J. P., Cantalupo, C. & Hopkins, W. D. 2006 Gesture handedness predicts asymmetry in the chimpanzee inferior frontal gyrus. *NeuroReport* **17**, 923–927. (doi:10.1097/01.wnr.0000221835.26093.5e)

46 Hopkins, W. D., Russell, J. L., Freeman, H., Buehler, N., Reynolds, E. & Schapiro, S. J. 2005 The distribution and development of handedness for manual gestures in captive chimpanzees (*Pan troglodytes*). *Psychol. Sci.* **16**, 487–493.

47 Hopkins, W. D., Russell, J. L., Hook, M., Braccini, S. & Schapiro, S. J. 2005 Simple reaching is not so simple: association between hand use and grip preferences in captive chimpanzees. *Int. J. Primatol.* **26**, 259–277. (doi:10.1007/s10764-005-2924-y)

48 Hopkins, W. D., Russell, J. L., Schaeffer, J. A., Gardner, M. & Schapiro, S. J. 2009 Handedness for tool use in captive chimpanzees (*Pan troglodytes*): sex differences, performance, heritability and comparion to the wild. *Behaviour* **146**, 1463–1483. (doi:10.1163/156853909X441005)

49 Hopkins, W. D., Wesley, M. J., Izard, M. K., Hook, M. & Schapiro, S. J. 2004 Chimpanzees are predominantly right-handed: replication in three colonies of apes. *Behav. Neurosci.* **118**, 659–663. (doi:10.1037/0735-7044.118.3.659)

50 Fragaszy, D., Izar, P., Visalberghi, E., Ottoni, E. B. & Oliveira, M. G. 2004 Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *Am. J. Primatol.* **64**, 359–366. (doi:10.1002/ajp.20085)

51 Moura, de. A. A. C. & Lee, P. C. 2004 Capuchin stone tool use in caatinga dry forest. *Science* **306**, 1909. (doi:10.1126/science.1102558)

52 Phillips, K. A. 1998 Tool use in wild capuchin monkeys (*Cebus albifrons trinitatis*). *Am. J. Primatol.* **46**, 259–261. (doi:10.1002/(SICI)1098-2345(1998)46:3<259::AID-AJP6>3.0.CO;2-R)

53 Westergaard, G. C. 1993 Hand preference in the use of tools by infant baboons (*Papio cynocephalus anubis*). *Percept. Motor Skills* **76**, 447–450. (doi:10.2466/pms.1993.76.2.447)

54 Westergaard, G. C. & Suomi, S. J. 1994 Hand preference in the use of nut-cracking tools by tufted capuchin monkeys (*Cebus apella*). *Folia Primatol.* **61**, 38–42. (doi:10.1159/000156726)

55 Westergaard, G. C. & Suomi, S. J. 1994 The use of probing tools by capuchin monkeys: evidence of increases right hand use with age. *Int. J. Primatol.* **15**, 521–529.(doi:10.1007/BF02735969)

56 Westergaard, G. C. & Suomi, S. J. 1996 Hand preference for stone artefact production and tool use by monkeys: possible implications for the evolution of right-handedness in hominins. *J. Hum. Evol.* **30**, 291–298. (doi:10.1006/jhev.1996.0024)

57 Mitchell, R. W. & Anderson, J. R. 1997 Pointing, withholding information, and deception in capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* **111**, 351–361.(doi:10.1037/0735-7036.111.4.351)

58 Li, L., Preuss, T. M., Rilling, J. K., Hopkins, W. D., Glasser, M. F., Kumar, B., Nana, R., Zhang, X. & Hu, X. 2009 Chimpanzee (*Pan troglodytes*) pre-central corticospinal system asymmetry and handedness: a diffusion magnetic reonance imaging study. *PLoS ONE* **5**, e12886. (doi:10.1371/journal.pone.0012886)

59 Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhang, X., Zhao, T., Hu, X. & Behrens, T. 2008 The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* **11**, 426–428. (doi:10.1038/nn2072)

Research

Comparative investigations of manual action representations: evidence that chimpanzees represent the costs of potential future actions involving tools

Scott H. Frey[1] and Daniel J. Povinelli[2],*

[1]*Department of Psychological Sciences, University of Missouri, Columbia, MO 65211, USA*
[2]*Department of Biology, University of Louisiana, Lafayette, LA 70504, USA*

The ability to adjust one's ongoing actions in the anticipation of forthcoming task demands is considered as strong evidence for the existence of internal action representations. Studies of action selection in tool use reveal that the behaviours that we choose in the present moment differ depending on what we *intend to do next*. Further, they point to a specialized role for mechanisms within the human cerebellum and dominant left cerebral hemisphere in representing the likely *sensory costs* of intended future actions. Recently, the question of whether similar mechanisms exist in other primates has received growing, but still limited, attention. Here, we present data that bear on this issue from a species that is a natural user of tools, our nearest living relative, the chimpanzee. In experiment 1, a subset of chimpanzees showed a non-significant tendency for their grip preferences to be affected by anticipation of the demands associated with bringing a tool's baited end to their mouths. In experiment 2, chimpanzees' initial grip preferences were consistently affected by anticipation of the forthcoming movements in a task that involves using a tool to extract a food reward. The partial discrepancy between the results of these two studies is attributed to the ability to accurately represent differences between the motor costs associated with executing the two response alternatives available within each task. These findings suggest that chimpanzees are capable of accurately representing the costs of intended future actions, and using those predictions to select movements in the present even in the context of externally directed tool use.

Keywords: motor planning; action selection; chimpanzee tool use; context sensitivity; cerebral asymmetry

1. INTRODUCTION

In the late nineteenth century, Donders [1] conducted a pioneering experiment in which he contrasted the times required to execute a simple (respond to the appearance of a stimulus) versus choice (select one among two responses based on the identity of the stimulus) response. The difference in response latencies between these conditions was interpreted as reflecting the additional time required for the cognitive process of selection. One of the many things that we have learned in the intervening century is that, independent of the hand used, the human left cerebral hemisphere plays a dominant role in this fundamental process [2]. In right-handed adults (some 90% of the population), increased activity is detected in the left posterior parietal cortex (along the intraparietal sulcus, IPS), inferior frontal sulcus extending into the rostral middle frontal gyrus (rMFG) and dorsal premotor cortex (dPMC) [3]. This network of regions may form a core system for response selection.

(a) *Response selection versus action selection*

In the typical response-selection task, the mapping between sensory stimuli and motor responses is both fixed and explicitly known by the actor (e.g. press the left key when the light is blue, or the right key when the light is red). This differs critically from the demands that typify real-world action selection, where there are often numerous potential responses (movements) that could be used to solve the problem at hand [4]. For instance, consider the range of hand postures that might suffice to grasp a mug of coffee and bring it to one's mouth for a drink. While much remains to be learned about how this degrees-of-freedom problem is solved, it is generally accepted that action selection is informed by predictions of the motor costs that would accompany various response options. To the extent that these forecasts are accurate, they enable us to select actions that lead to successful (rewarding) solutions to the problem while minimizing costs [5].

* Author for correspondence (povinelli@louisiana.edu).

(b) Neural substrates of response and action selection

In order to gain further insights into the role of prediction in action selection, we have used a simplified task in which participants are required to choose whether to engage an object (e.g. a handle) appearing in different orientations in an under- versus over-hand grip (e.g. a power grip). In all trials of these experiments either grip is physically possible, the question is which is preferable. We consistently find that participants prefer to grasp stimulus objects using the alternative that is perceived as least costly (or awkward) during overt execution. This is even true when they are asked to make their choices under prospective grip-selection (PGS) conditions, where movements are never actually undertaken [6]. The degree of correspondence between grip preferences in overt and PGS tasks suggests that even in the complete absence of feedback, participants are able to predict the likely motor costs of future actions with high fidelity, and select actions accordingly. Furthermore, this ability seems to be retained by many patients even during periods of acute [7] or chronic [8,9] limb disuse, or following amputation [10].

Our early functional magnetic resonance imaging work on prospective power grip selection revealed increases in the dPMC, superior parietal lobule (SPL) and along the IPS [11]. More recent findings show that prospective precision grip selection based on either hand engages the entire core network of regions implicated in response selection (left IPS, rMFG and dPMC), as introduced earlier. In addition, we find increases within a number of other brain regions including the bilateral cerebellum, SPL, pre-supplementary motor area, right dPMC, as well as in the left anterior intraparietal sulcus (aIPS) and left ventral premotor cortex (vPMC) [12]. After physical practice, these same areas come to represent PGS decisions based on the use of a formerly novel grasping tool that differs mechanically and dynamically from the natural limbs.

One possible interpretation is that these regions participate in estimating the motor costs associated with the two grip alternatives: under- or over-hand. An emerging view is that the cerebellum supports forward internal models that predict the likely sensory consequences of a motor command slightly in advance of the actual sensory feedback that accompanies movement [13,14]. These feed-forward predictions are thought to update multi-sensory estimates of the state of the body (e.g. posture of the upper limbs), represented in the parietal cortex (particularly the SPL) [13,15]. An interesting possibility is that these same predictive mechanisms might play a role in forecasting the long-range consequences of response alternatives [5,16,17]. This information could be valuable to action selection by providing a means of estimating the motor costs (energy expenditure, awkwardness) of candidate responses and their potential to achieve the desired reward state, representations that may be computed subcortically in the basal ganglia and/or brain stem [18].

(c) Cerebral asymmetry

The left cerebral asymmetry for PGS responses in vPMC and the aIPS is striking in comparison with the bilateral engagement of dPMC and the SPL. In humans, left vPMC is situated in the inferior portion of the precentral gyrus located immediately caudal to the *pars opercularis* (or Broca's area), which is known to be involved in a variety of language-, and a growing list of action-related functions [19,20]. In monkeys, rostral ventral premotor cortex (F5) is the putative homologue for *pars opercularis* [21], and is implicated in a variety of higher level motor functions including the representation of action goals [22]. Caudal ventral premotor cortex (F4), by contrast, is implicated in constructing multi-sensory representations of peripersonal space and of limb and head movements [23]. Our present understanding of vPMC functions in the human brain is limited. Though highly speculative, it is possible that with the emergence of language functions in rostral vPMC, the human left caudal vPMC has become more heavily involved in the representations of action goals. Some work demonstrating increased activity in this region during action perception appears consistent with this view [24].

There is mounting evidence for the role of the human aIPS in action representation [25,26], and more than a century of data exist implicating the left parietal and frontal cortex in manual praxis [27,28]. Asymmetrical involvement of the left aIPS in these planning tasks may be related to the fact that this region has direct anatomical connections with the vPMC [29], allowing these two areas to operate as a functional unit. Following the line of reasoning above, it may be that the left parietal asymmetry is a consequence of changes in the role of caudal vPMC precipitated by the emergence of language in rostral vPMC. This is highly speculative; however, recent findings do suggest that cerebral asymmetries in posterior parietal action representations (Brodmann area 40) are correlated with lateralization of language in Broca's area (Brodmann area 44/45) and its right hemisphere homologue [30].

(d) Context sensitivity in action selection

In speech, it is well known that articulation of a phoneme is affected by the identity of upcoming phonemes. Analogous effects of task context have been reported in a wide variety of manual behaviours including: typing, handwriting, manual aiming and prehension (see review in Johnson-Frey *et al.* [31]). Action-selection tasks have proven to be especially valuable in studying context effects [32], and have shed light on the properties of underlying movement representations [33]. Grip-selection tasks involving tools have proven to be particularly useful in revealing details of motor planning in human infants and adults [34]. Context effects in manual action selection have been shown to emerge during the first 2 years of life [35–38], and development can be accelerated with training [39]. While the physical properties of the effectors may contribute to some context effects [40], many of these findings are difficult to interpret without acknowledging a role for internal representations of task demands that go beyond immediately available sensory information [41,42]. This point is critical to understanding behaviours where the choice

of a response can be influenced by the goal of the larger action sequence in which it is embedded. As a consequence, the responses that we choose in the present moment differ depending on how they might impact the costs of what we *intend to do next*.

(e) *Evidence for predictive action selection in non-human primates*

Whether similar predictive mechanisms exist in other species is an important and challenging question. A first step is to ask whether evidence can be found for context sensitivity in action-selection behaviours similar to those identified in humans. Several studies have tested whether, like humans, monkeys show an *end-state comfort effect*, i.e. whether they adapt their initial grip of an object in a way that reflects anticipation of movements required to achieve the subsequent goal of the task [43]. Across two foraging tasks, Weiss *et al.* [44] provided evidence for end-state comfort effects in the grip preferences of cotton-top tamarins. They later extended this finding to include lemurs, a group of primates even more distantly related to humans [45]. More recently, Nelson *et al.* [46] demonstrated that after a limited number of trials, most of the rhesus macaques they tested were able to develop a grip to bring a baited spoon to their mouths in an efficient manner (radial grip with the thumbside of the hand towards the bowl of the spoon) across 12 trials. An important issue is whether these behaviours reflect the learning of stimulus-response contingencies, or involve anticipation of motor costs. Evidence shows that this efficient grip was learned during the course of testing. This can be derived from the fact that, on difficult trials, the monkeys' performances improved from 28 per cent efficient grips in the first session of testing to 94 per cent in the second session. Of the six monkeys tested, three switched hands depending on the side towards which the bowl of the spoon was oriented; the remaining three used the same hand (and grip), but changed their body position relative to the spoon. No monkeys switched from an over-hand grip on the easy trials to an under-hand grip on the difficult trials. After this rapid acquisition, the monkeys maintained these efficient strategies even 1 year later when they were retested.

Given that several disparately related primate species have shown some evidence for anticipatory planning in action-selection tasks (i.e. the end-state comfort effect), it is tempting to conclude that this is a quite anciently evolved ability widely present in non-human primates, and perhaps even other mammals. While prospective planning abilities might be critical for tool-use behaviours, they would seem to have evolved in species that are not known to use tools in nature [45]. However, the existing evidence is currently limited to situations in which a simple tool is directed to the subject's mouth. In humans, anticipatory planning in the context of self-directed tool abilities emerges during the first 2 years of life [38]. However, the development of planning in externally directed tool-use tasks is more difficult and lags behind even when feedback is enhanced [37]. It is not known whether non-human primates show this more advanced form of action planning. For example, do they exhibit the end-state comfort effect when grasping a tool and directing it towards a goal located in extra-personal space?

Here, we report findings from studies (conducted between 2001 and 2004) that we designed and conducted to explore this question in our nearest living relatives, chimpanzees. Unlike lemurs, tamarins or rhesus macaques, chimpanzees both use and make tools as part of their natural ecology, and the particular individuals we tested had a long history of doing so in the laboratory in a wide variety of tasks. Like humans, chimpanzees regularly perform actions with tools that are directed towards the self (e.g. eating termites from a stick) and towards external targets (e.g. inserting a stick into a termite nest). Our strategy was to use an analogue of the procedure used by McCarty *et al.* [38] to study the development of the end-state comfort effect in spoon-feeding in toddlers (experiment 1), and to then use these results as a platform to study the more advanced abilities involved in using a tool to obtain a food reward in extra-personal space (experiment 2).

2. EXPERIMENT 1: GRIP SELECTION IN A SELF-DIRECTED TOOL-USE TASK IN CHIMPANZEES

In our initial investigation, we sought to identify whether our chimpanzees would display an end-state comfort effect in their grip selection during a self-directed feeding task with similarities to the one used to test macaques following the study of Nelson *et al.* [46]. The stimulus was a horizontal tool (dowel) with either the left- or right-end baited. If action selection is influenced by the intended subsequent movement (bringing the baited end of the handle to the mouth), then we reasoned that subjects would prefer the grip that places the thumbside of the hand towards the baited end; i.e. they would prefer a radial grip. Choosing the alternative ulnar grip would be a more costly option, making it more difficult to bring the food to the mouth. Grip preferences in this test condition were compared with the control in which the choice of grip was irrelevant because both ends of the tool were baited.

(a) *Method*
(i) *Subjects*
Five adult female chimpanzees (age range = 13 years, seven months to 14 years, six months) participated in the study. The subjects were housed at the University of Louisiana and had participated in numerous dowel-use studies over a period of over 10 years [47].

(ii) *Apparatus*
A 30 cm long dowel (plastic pipe) was used in the experiment. In the experimental conditions, the dowel rested horizontally on two L-shaped brackets (approx. 20 cm apart and 50 cm above the floor) that were attached to the wall of the subjects' testing unit. One or both ends of the dowel could be easily covered with a highly desirable food reward (such as peanut butter or honey).

(b) Procedure

(i) Orientation and food preference

We placed 14 (unbaited) replicas of the dowel in the subjects' indoor–outdoor living environment. This allowed the chimpanzees to interact and familiarize themselves with the tool. If the apes threw the dowels out of the enclosure, then they were returned by the caretakers. We made these replicas available to the apes in their living environment every day throughout all phases of the experiment.

One week after we introduced the dowels, the primary trainer and the caretaker (hereafter referred to as the trainer) individually brought each ape into an indoor testing unit. The test unit was connected to an outdoor waiting area by a shuttle door that could be remotely opened and closed to allow the ape the opportunity to enter and exit the test unit. The apes were highly familiar with this procedure (see Povinelli [47]). The trainer inserted two dowels into the test unit and held them as the ape approached. Each dowel was baited with a different reward. The first dowel chosen to eat from was recorded. Each ape was administered 10 trials. The first reward chosen seven or more times out of 10 was considered their preferred reward. Any ape not exhibiting a preference was given both rewards randomly and equally across conditions throughout the study.

(ii) Test orientation

We conducted a series of unstructured three-trial sessions in which the dowel was placed on the bracket with food baited on both ends. The caretaker allowed each ape to enter the test unit individually, or in pairs, until such a time as they reliably took the dowel from the bracket and consumed the food. The shuttle door was opened as soon as the apes had finished eating the bait and dropped the dowel, or after 1 min from the time they entered the test unit. The apes then began testing sessions.

(iii) Testing

Testing took place in the test unit and consisted of 60 trials per ape, with a maximum of six trials per day. Before each trial began, the trainer baited a dowel on both ends and placed it horizontally on the brackets on the wall to the ape's left as he or she entered. The ape then entered the testing unit and the shuttle door was closed. After the ape grasped the dowel and consumed the reward, the trainer opened the shuttle door and allowed the ape to exit.

Testing consisted of two types of trials: (i) on test trials only, one end of the dowel was baited (on half of the test trials the right end was baited and on the other half the left end was baited) and (ii) on control trials both ends of the dowel were baited. Apes were given 40 test trials (20 baited on the right end and 20 baited on the left end) and 20 control trials. The trials were randomly administered within the following constraints: (i) the same type of trial was not administered on more than three sequential trials, (ii) the numbers of each type of trial were equal for the first and second halves of the study, and (iii) the left/right
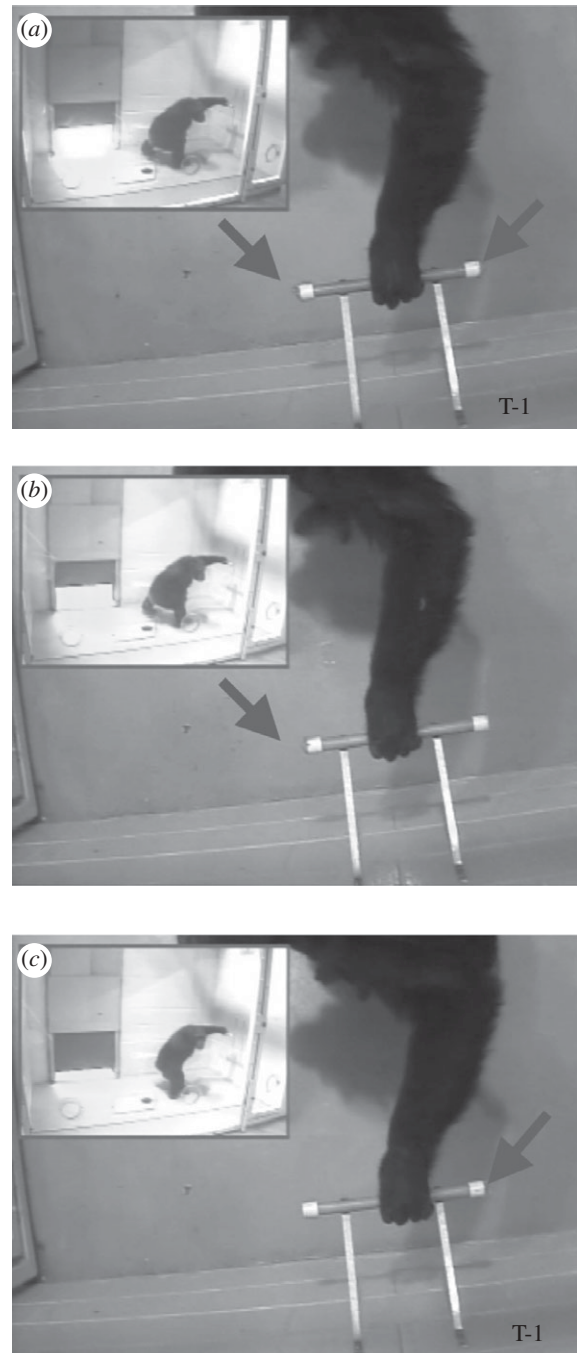
Figure 1. The subject Brandy grasps the dowel in the three experimental conditions of experiment 1: (*a*) both-ends baited, (*b*) right-end baited, and (*c*) left-end baited. Note that her grip selection is the same in all three cases. Arrows indicate position of the bait (peanut butter).

orientations of the dowels on the test trials were not the same for more than two consecutive trials.

One experimenter was present in the rear of the test unit to control the shuttle door. Every trial was recorded on video with a view that allowed excellent visibility regarding which hand the ape used to grasp the dowel, the position of the hand while reaching for the pipe and the type of grip used (figure 1).

(iv) Coding and inter-rater reliability

A main rater scored all trials for the orientation of the subject's thumb (up or down). This allowed for an
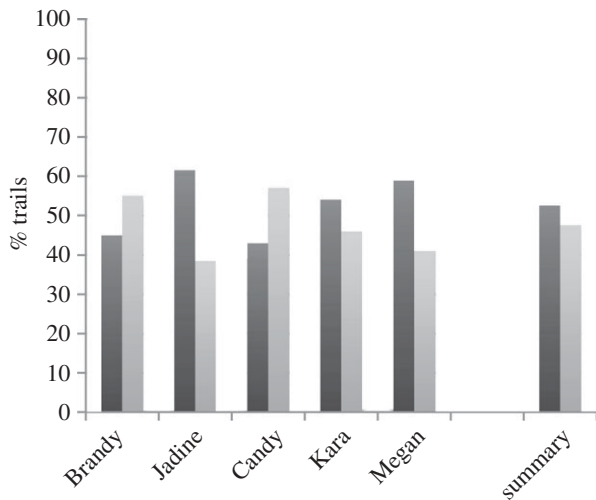
Figure 2. Grip preferences in experiment 1. Three of the participants (Jadine, Kara and Megan) exhibit a trend towards preferring to grasp the handle with the thumb towards the baited end. This is consistent with what would be expected if the animals were taking the desired end state into consideration when selecting their grasps. These trends were not statistically reliable (see text for details; black bars, thumb towards bait; grey bars, thumb away from bait).

Table 1. Summary of thumb placement (in per cent of trials) as a function of experimental condition.

| | conditions | | |
thumb placement	control	baited-left	baited-right
Brandy			
left	10	5	15
right	90	95	85
Jadine[a]			
left	75	95	78
right	25	5	22
Candy[b]			
left	89	79	87
right	11	21	13
Kara			
left	100	100	95
right	0	0	5
Megan[c]			
left	90	100	84
right	10	0	16

[a]One trial was inadvertently not recorded.
[b]An experimental error resulted in the administration of 18 control trials and 19 baited left and 23 trials baited right test trials.
[c]One trial was not codable owing to the obstruction of the stimuli by the subject's body.

unambiguous determination of whether their grip changed as a function of trial type (i.e. from overhand to under-hand) as a function of how the dowel was baited (right end, left end and both ends). A secondary rater examined 25 per cent of the trials and agreed with the main rater on 96 per cent of all trials.

(c) Results and discussion

The individual apes grasped the dowel in a variety of ways (over-hand, under-hand, pincer grip, between fingers). However, each ape exhibited a striking consistency in how they did so regardless of the experimental condition. For example, figure 1a–c displays the ape Brandy grasping the dowel across the three-trial types: (i) both ends baited, (ii) right-end baited, and (iii) left-end baited. Notably, she uses the same hand and grip in each condition. Although all apes exhibited some variation, the position of their thumb upon initial grasp (towards the baited end or away from the baited end) provided a highly reliable measure of whether they switched grips as a function of what end of the dowel was baited on the test trials. Figure 2 displays the thumb position (towards or away from the baited end) for each hand (right and left). In seeming contrast to what would be expected if the apes were selecting their actions based on prediction of the motor costs that would be experienced, as a group the apes did not strongly alter their grip as a function of which end was baited. Instead, they simply grasped the dowel in a habitual fashion and inserted the baited end into their mouths. The results of the control trials confirm that the apes had a habitual grip preference: they did not depart from what they displayed in the test trials (table 1). In other words, we saw no evidence that action selection is context-sensitive.

Three of the five apes, however, showed a preference for the thumb towards the baited end of the dowel (radial grip), and one (Jadine) showed a fairly strong bias for the thumb towards the baited end (61.5% or 24/39 of trials; table 1). Interestingly, Jadine is the one ape that showed fairly quick improvement in a previous test, wherein the apes needed to learn to grasp a tool by a non-functional end in order to use the other functional end (see experiment 2 below; [47], experiment 13). Megan also exhibited a similar trend in favour of placing the thumb towards the baited end (59% or 23/39 of the trials; table 1). However, binomial tests (one-tailed, chance = 0.5) indicated that none of the animals (including Jadine and Megan) exhibited a statistically reliable difference between these two grip options ($p < 0.09$ or smaller in all cases).

These results would appear to indicate that chimpanzees were not representing forthcoming task demands and modifying their initial grip preferences accordingly. None of the individuals showed significant differences between grip preferences in the test condition and those exhibited in the control. This differs dramatically from results demonstrated by toddlers when grasping wooden spoons for self-feeding [38] and previous investigations with monkeys discussed earlier [44–47]. It is tempting to conclude that these apes lack the ability to represent the costs of forthcoming task demands and/or to adapt their responses accordingly even in a self-directed tool-use task. Indeed, in the test trials, the apes frequently wound up with the dowel in what appeared to be an awkward position. For example, an over-hand grasp (palm down) with the thumb oriented away from the baited end of the dowel left the apes in a biomechanically awkward position when bringing the tool to the mouth. However, the apes compensated for this by further rotating their wrists, and/or by tilting their heads. Although this appeared awkward to human

Figure 3. In experiment 1, the subject Brandy has just used an over-hand grip (palm down) with thumb pointing away from the baited end of the dowel (figure 1*c*). As a consequence, she must lift the dowel higher and rotate the baited end towards her mouth. This involves considerably more effort than the movement that results from the thumb towards the baited end.

observers, nevertheless, the apes, were adept at eating the food off the baited end of the dowel in this manner (figure 3). What we can say with confidence is that in such cases, the ape was required to exert greater motor effort in order to rotate the baited end a longer distance to reach their mouths. Nonetheless, any additional costs associated with ending the handle rotation in an awkward posture may not have been sufficient to result in anticipatory modification of the initial grip. This issue arises in grip-selection studies involving humans where participants show stronger end-state effects when the levels of accuracy for final object positioning are increased [48]. This possibility is addressed in our second experiment where the costs of choosing the incorrect initial grip would be maximal: failure to complete the trial and obtain a food reward. If we are correct, then the apes should exhibit evidence for anticipatory grip selection in this more challenging task.

3. EXPERIMENT 2: GRIP SELECTION IN A TOOL-USE TASK DIRECTED TOWARDS AN EXTERNAL TARGET IN CHIMPANZEES

In this study, chimpanzees needed to use a thick dowel to dislodge and obtain a food reward. A critical feature is that the task could only be completed if the animals grasped the dowel with the thumbside of their hand towards its centre. We tested for anticipatory effects by analysing how the apes gripped the dowel when presented in a variety of different orientations.

(a) Method
(i) Subjects
The five chimpanzees from experiment 1, plus two other adult members of their group, participated in the study. At the time the study began, the apes ranged in age from 15 years, four months to 16 years, three months. The apes had participated in numerous studies similar to the one used here (i.e. using a dowel or stick to dislodge a reward from a platform;
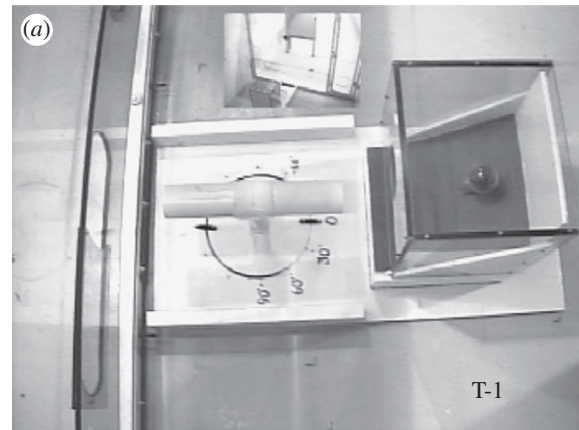
Figure 4. Experiment 2. (*a*) The apparatus with the dowel in the 0° orientation. This set-up was used in the demonstration and criterion phase, as well as during the 0° trials of the testing phase. (*b*) An ape grasps the dowel in preparation to insert and dislodge the goal/reward.

see Povinelli [47]). However, only one of these studies had systematically altered the orientation of a tool (see Povinelli [47], experiment 13). It is important to note that this previous study (and several related ones) was more cognitively demanding. It required the apes to understand that one end of the tool was functional and the other end was not, and to anticipate this in their initial grasp. In the present study, both ends of the tool were identical and equally functional.

(ii) Apparatus
The apparatus depicted in figure 4 was used. It consisted of a dowel that rested on a bracket that could be horizontally rotated, thus altering how the subject could grasp it. The goal of the task was to grasp the dowel, lift it off the bracket and poke it through a hole in the Plexiglass box to dislodge an apple (or some other round food reward). Once the ape tapped the apple with the dowel, it rolled to within their reach.

The dowel was 46 cm long and 8 cm in diameter. In order to deter apes from gripping the centre of the dowel, a 10 cm wide Velcro strip was wrapped around the dowel's midpoint. (One ape was allowed a slightly modified dowel, owing to her aversion to the original stimulus. This dowel was 5 cm in diameter and instead of a Velcro strip, two black lines were painted on the dowel.)
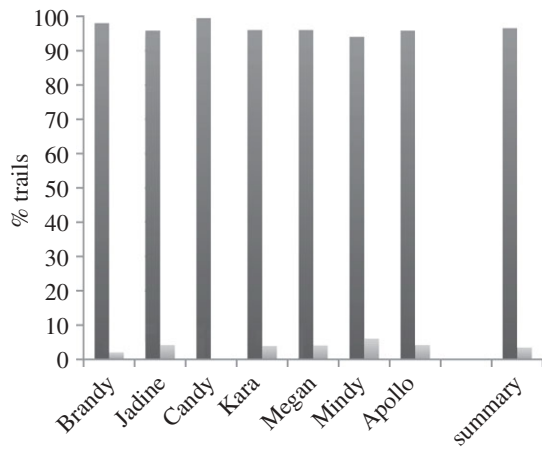
Figure 5. Grip preference results in experiment 2. All of the participants show some evidence of preferring to grasp the handle with the thumb towards the centre. This is consistent with what would be expected if the animals were taking the desired end state into consideration when selecting their grasps (black bars, thumb towards centre; grey bars, thumb away from centre).

(b) Procedure

(i) Familiarization with the dowel

In several sessions in the test unit, each ape was exposed to dowels of varying thicknesses, including ones with Velcro around the centre. No apparatus was present and the apes did not use the dowels to obtain food of any kind. Rather, these sessions simply helped us to identify the maximum thickness the dowel could be and still allow the apes to easily grasp it, and to ensure that the Velcro strip deterred apes from grasping the dowel in the centre.

(ii) Demonstration of the task

The trainer brought each ape into the test unit individually for a demonstration session in which the experimenter demonstrated how to use the dowel to dislodge the food reward from the apparatus. First, the apparatus (figure 4) was positioned on the trainer's side of the Lexan partition with the dowel placed on the bracket in a 0° horizontal orientation (i.e. perpendicular to the apparatus; figure 4). The apparatus was thus positioned directly in front of an opening in the Lexan partition through which the chimpanzees could reach and grasp the dowel. Next, the trainer opened the shuttle door and the ape entered the test unit. The trainer then closed the shuttle door and demonstrated the task by picking up the dowel and dislodging the food by poking it through the hole in the apparatus. The trainer then handed the food to the ape through the hole in the Lexan door. Finally, the trainer opened the shuttle door and ushered the ape out of the test unit.

(iii) Criterion

The criterion consisted of two-trial sessions that immediately followed the demonstration session. The trainer positioned the apparatus directly in front of the opening in the Lexan partition and the dowel was positioned on the bracket directly in front of the response slot (centred in front of the opening in the

apparatus) on the experimenter's side of the Lexan. For all criterion trials, the dowel was oriented at 0° (figure 4). After the shuttle door was opened, the ape was given 1 min to enter the test unit and an additional maximum of 1 min to initiate the trial by touching a ready-to-respond (RTR) symbol. This resulted in the hole in the partition opening, thus allowing the ape to reach out and grasp the dowel. A complete response was defined as the ape touching the RTR symbol, grasping the dowel and inserting it through the opening in the apparatus and dislodging the food reward so that it fell down the ramp towards the ape. If the ape did not touch the RTR symbol and/ or touch the dowel, then the trial was scored as no response and was immediately re-run. If the ape initiated a trial through the RTR procedure and grasped the dowel but did not complete the response, then the trial was scored as incomplete and not re-run. If any of the following conditions were met, then the response barrier was raised immediately, the ape was ushered out of the test unit, and an incomplete response was recorded: (i) the ape initially lifted the dowel off the bracket using two hands, (ii) the ape displaced the bracket from its place in the spacer board, (iii) the ape touched the apparatus with their hand or arm prior to completing the task, and/or (iv) the dowel fell out of the ape's reach before completing the task.

Any trials scored as incomplete were re-conducted at the end of all testing sessions for all apes. The participants were required to finish four complete trials across two consecutive sessions in order to move into testing.[1]

(iv) Testing

Testing consisted of 24 two-trial sessions per ape, for a total of 48 test trials (eight for each of six dowel orientations; see below). The configuration of the test unit and procedure remained the same as in the criterion phase except that the dowel was presented in one of the six orientations. Before each trial, the dowel was oriented at 0°, 30°, 60°, 90°, −30° or −60° (with 0° being perpendicular to the Lexan (one end pointing towards the ape and the other end pointing away), 30° being a clockwise rotation, and 90° being parallel to the Lexan partition). Each of the orientations was presented in a randomized order with the constraint that each orientation occurred once before any repeat, and all orientations occurred four times within the first and last half of the total number of trials. A complete response was defined as the ape touching the RTR symbol, grasping the dowel and contacting the dowel to the Plexiglass panel on the front of the apparatus and/or inserting the dowel through the hole in the apparatus. All trials scored as incomplete were re-run immediately. A camera was mounted directly above the bracket and dowel in order to obtain a clear view of the ape's grip.

(v) Coding and inter-rater reliability

All trials were coded for the position of the ape's thumb: either towards the centre of the dowel or away from the centre of the dowel. Critically, because of the thickness of the dowel and the biomechanical
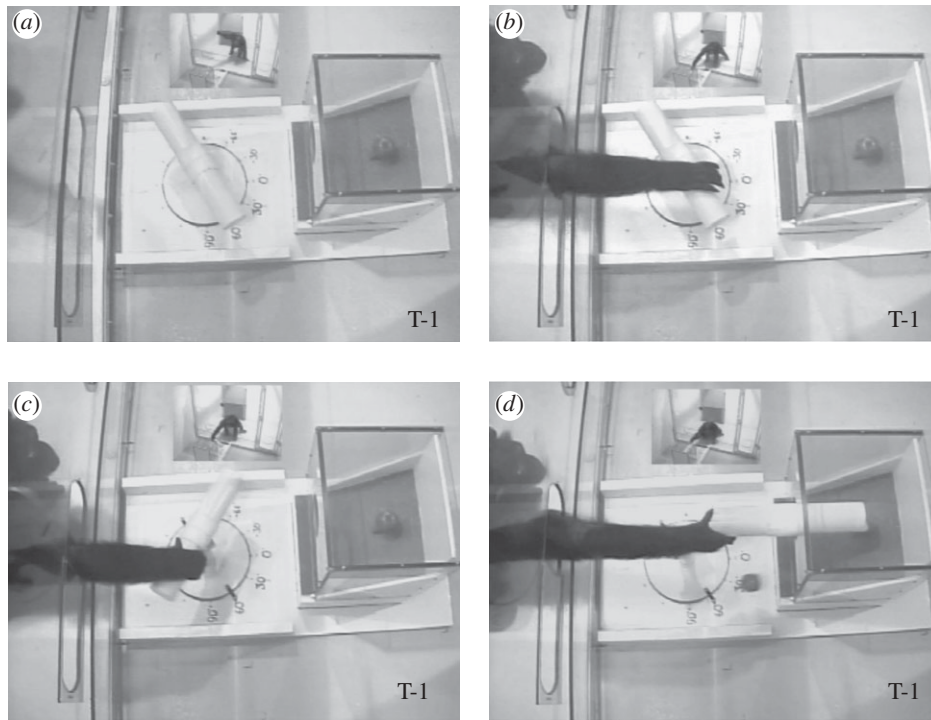
Figure 6. An illustration of the anticipatory grip-selection effect in the externally directed tool-use task: (*a*) the subject Kara reaches out and selects the end of the dowel that is farthest away so as to have her thumb positioned towards the centre of the dowel. (*b*–*d*) This allows the tip of the dowel to swing into a position aligned with the hole, thus allowing for easy insertion and the dislodging of the apple.

constraints of reaching through the Lexan partition and orienting the dowel into the hole in the apparatus, the apes could only dislodge the apple if they grasped the tool with the thumb pointed towards the centre. Two raters independently scored the video recordings for thumb directions and agreed on 99 per cent (332/335) of all trials (one trial was inadvertently not recorded).

(c) *Results and discussion*
As is clear in figure 5, each of the apes displayed evidence of anticipatory grip selection in this task. All individuals exhibited a strong preference for grasping the dowel with the thumbside of their hands towards the centre of the dowel across orientations. Binomial tests (one-tailed, chance $= 0.5$) for each animal indicated that all of the apes exhibited a statistically reliable effect of grasping the dowel with thumb towards the centre of the dowel ($p < 0.001$ in all cases). These results demonstrate that as the dowel orientation was manipulated, the subjects switched either the side of the dowel they grasped (left or right) and/or their hand orientation (palm up versus down) in order to keep their thumb oriented towards the centre of the dowel. If they had not performed so, then they would have frequently grasped the dowel in a manner that made it biomechanically impossible for them to insert the dowel through the hole and dislodge the apple (without first stopping to adopt a new grip). Thus, in stark contrast to the results of experiment 1, these data provide evidence that chimpanzees are capable of anticipatory grip selection in a tool-use task. Because the apes were at ceiling levels across all dowel

orientations, our results also highlight the fact that the apes did not learn to modify their grips in this manner across trials. Rather, they exhibited immediate evidence of grasping the dowel in the way that would allow for successful manipulation and dislodging of the apple.

As an example of the effect, consider figure 6*a*–*d*. Here, the subject Kara reaches out to grasp the right end of the dowel with her right hand (thumb pointed towards the centre of the dowel), even though the right end of the dowel is further away from her than the left end. This allows her to swing the free end of the dowel towards the hole through which the dowel needs to be inserted. If she had instead grasped the left side of the dowel with her right hand (thumb pointed away from centre), then Kara would have been unable to insert the free end into the Plexiglass box and dislodge the apple. The facts that the dowel was rotated 180° across trials in a stepped fashion, and that the apes virtually always grasped the dowel with the thumb pointing towards the centre, show that they selected the grip best suited to achieving the action goal before grasping the dowel.

In sum, the results of experiment 2 establish that, in the context of a tool-using task, chimpanzees select from among two actions the one that will be successful. Importantly, this behaviour does not seem to have been learned over many trials. Instead, the apes' grip preferences appear to involve anticipatory planning. We hypothesize that they selected their responses through the formation of internal representations of the motor costs and probability of reward associated with the available response options. If this is correct, however, then why were their grip preferences in experiment 1, an

ostensibly less difficult, self-directed task, not significantly affected by changes in the task context? The reason may be that the anticipated motor costs of obtaining the reward by placing the thumb towards or away from the baited end of the tool were simply too similar and negligible to influence grip selection. Individual apes thus chose one approach and generally stuck with it across all conditions.

4. CONCLUSIONS

Action selection is one of the fundamental problems that must be solved for behaviour to be adaptive ([49] and associated articles). There is mounting evidence that a key function of the extended motor system is generating predictions of the sensory consequences that are likely to arise from future movements, and using that information to estimate motor costs and select actions. Investigations of the grip-selection behaviours of a variety of primates suggest that these feed-forward processes are not unique to humans, and may exist in species that are not known users of tools in nature, including lemurs [45], tamarins [44] and rhesus monkeys [47]. Our findings with chimpanzees extend these previous results by demonstrating that some non-human primates are able to cope with action-selection planning even when tools are used to interact with goal objects located distally (away from the body).

Whether the neural mechanisms involved in anticipatory planning are organized similarly across primate species remains unknown. As reviewed earlier, there is mounting evidence from humans for the involvement of cerebellar, cortical and subcortical mechanisms in these predictive functions, with inferior parietal and ventral premotor cortices showing a pronounced left cerebral asymmetry. Similar to the tasks involved here, recent findings suggest that activity within these very same regions is increased when grip selection involves representing forthcoming task demands, e.g. when selecting how best to grasp a handle for the purpose of performing a subsequent rotation [50]. Whether this left-cerebral asymmetry is causally related to right-hand dominance and/or the evolution of language is unknown. However, the left-cerebral asymmetry for action planning may be coupled with right-hand dominance. Strongly left-handed individuals (who show greater incidence of atypical language organization) display increased activity in both left and right vPMC during PGS tasks involving the hands or a tool [51]. While it is true that non-human primates demonstrate hand preferences under certain circumstances [52,53], these never approach the population-level right-hand bias evident in roughly 90 per cent of humans for fine motor tasks [54,55]. The fossil and archaeological records suggest that this right-hand bias was evident very early in our lineage [56]. The emerging picture is that predictive planning may extend much further back into our primate origins, and is likely rooted in basic functions of sensorimotor control that predate handedness. The ability to predict the future based on past experiences is a core cognitive faculty of modern humans [57], and may have played a critical role in the evolution of more sophisticated forms of tool manufacture and use [58].

ENDNOTE

[1]Additional demonstration sessions were conducted for any ape that did not appear to understand the task, until it demonstrated such an understanding. These apes were given remedial training after three unsuccessful demonstrations and criterion trials (see §3c).

REFERENCES

1 Donders, F. C. 1969 On the speed of mental processes. *Acta Psychol.* **30**, 412–431. (doi:10.1016/0001-6918(69)90065-1)
2 Rushworth, M. F., Nixon, P. D., Wade, D. T., Renowden, S. & Passingham, R. E. 1998 The left hemisphere and the selection of learned actions. *Neuropsychologia* **36**, 11–24. (doi:10.1016/S0028-3932(97)00101-2)
3 Schluter, N. D., Krams, M., Rushworth, M. F. & Passingham, R. E. 2001 Cerebral dominance for action in the human brain: the selection of actions. *Neuropsychologia* **39**, 105–113. (doi:10.1016/S0028-3932(00)00105-6)
4 Bernstein, N. 1967 *The co-ordination and regulation of movements.* Oxford, UK: Pergamon Press.
5 Shadmehr, R. & Krakauer, J. 2008 A computational neuroanatomy for motor control. *Exp. Brain Res.* **185**, 359–381. (doi:10.1007/s00221-008-1280-5)
6 Johnson, S. H. 2000 Thinking ahead: the case for motor imagery in prospective judgements of prehension. *Cognition* **74**, 33–70. (doi:10.1016/S0010-0277(99)00063-3)
7 Johnson, S. H. 2000 Imagining the impossible: intact motor representations in hemiplegics. *Neuroreport* **11**, 729–732. (doi:10.1097/00001756-200003200-00015)
8 Jenkinson, P. M., Edelstyn, N. M. & Ellis, S. J. 2009 Imagining the impossible: motor representations in anosognosia for hemiplegia. *Neuropsychologia* **47**, 481–488. (doi:10.1016/j.neuropsychologia.2008.10.004)
9 Johnson, S. H., Sprehn, G. & Saykin, A. J. 2002 Intact motor imagery in chronic upper limb hemiplegics: evidence for activity-independent action representations. *J. Cogn. Neurosci.* **14**, 841–852. (doi:10.1162/089892902760191072)
10 Philip, B. A. & Frey, S. H. 2010 *Intact prediction of grip selection with an amputated hand* (program no. 291.13, online). San Diego, CA: Society for Neuroscience.
11 Johnson, S. H., Rotte, M., Grafton, S. T., Hinrichs, H., Gazzaniga, M. S. & Heinze, H. J. 2002 Selective activation of a parietofrontal circuit during implicitly imagined prehension. *Neuroimage* **17**, 1693–1704. (doi:10.1006/nimg.2002.1265)
12 Jacobs, S., Danielmeier, C. & Frey, S. H. 2010 Human anterior intraparietal and ventral premotor cortices support representations of grasping with the hand or a novel tool. *J. Cogn. Neurosci.* **22**, 2594–2608. (doi:10.1162/jocn.2009.21372)
13 Wolpert, D. M. & Ghahramani, Z. 2000 Computational principles of movement neuroscience. *Nat. Neurosci.* **3**(Suppl. 11), 1212–1217. (doi:10.1038/81497)

14 Wolpert, D. M., Miall, R. C. & Kawato, M. 1998 Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–347. (doi:10.1016/S1364-6613(98)01221-2)

15 Wolpert, D. M., Goodbody, S. J. & Husain, M. 1998 Maintaining internal representations: the role of the human superior parietal lobe. *Nat Neurosci* **1**, 529–533. (doi:10.1038/2245)

16 Frey, S. H. 2010 Forecasting the long-range consequences of manual and tool use actions: neurophysiological, behavioral and computational considerations. In *Motor control: theories, experiments and applications* (eds F. Danion & M. Latash), pp. 295–313. Oxford, UK: Oxford University Press.

17 Grush, R. 2004 The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* **27**, 377–396 (discussion 396–442).

18 Humphries, M. D., Gurney, K. & Prescott, T. J. 2007 Is there a brainstem substrate for action selection? *Phil. Trans. R. Soc. B* **362**, 1627–1639. (doi:10.1098/rstb.2007.2057)

19 Binkofski, F. & Buccino, G. 2004 Motor functions of the Broca's region. *Brain Lang.* **89**, 362–369. (doi:10.1016/S0093-934X(03)00358-4)

20 Rizzolatti, G. & Craighero, L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)

21 Preuss, T. M., Stepniewska, I. & Kaas, J. H. 1996 Movement representation in the dorsal and ventral premotor areas of owl monkeys: a microstimulation study. *J. Comp. Neurol.* **371**, 649–676. (doi:10.1002/(SICI)1096-9861(19960805)371:4<649::AID-CNE12>3.0.CO;2-E)

22 Umilta, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. 2008 When pliers become fingers in the monkey motor system. *Proc. Natl Acad. Sci. USA* **105**, 2209–2213. (doi:10.1073/pnas.0705985105)

23 Rizzolatti, G., Fogassi, L. & Gallese, V. 2002 Motor and cognitive functions of the ventral premotor cortex. *Curr. Opin. Neurobiol.* **12**, 149–154. (doi:10.1016/S0959-4388(02)00308-2)

24 Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A. & Rizzolatti, G. 2004 Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. *J. Cogn. Neurosci.* **16**, 114–126. (doi:10.1162/089892904322755601)

25 Hamilton, A. F. & Grafton, S. T. 2006 Goal representation in human anterior intraparietal sulcus. *J. Neurosci.* **26**, 1133–1137. (doi:10.1523/JNEUROSCI.4551-05.2006)

26 Tunik, E., Rice, N. J., Hamilton, A. & Grafton, S. T. 2007 Beyond grasping: representation of action in human anterior intraparietal sulcus. *Neuroimage* **36**(Suppl. 2), T77–T86. (doi:10.1016/j.neuroimage.2007.03.026)

27 Heilman, K. M. & Rothi, L. J. G. 1997 Limb apraxia: a look back. In *Apraxia: the neuropsychology of action* (eds L. J. G. Rothi & K. M. Heilman), pp. 7–18. Hove, UK: Psychology Press/Erlbaum (UK) Taylor & Francis.

28 Leiguarda, R. C. & Marsden, C. D. 2000 Limb apraxias: higher-order disorders of sensorimotor integration. *Brain* **123**, 860–879. (doi:10.1093/brain/123.5.860)

29 Rushworth, M. F., Behrens, T. E. & Johansen-Berg, H. 2006 Connection patterns distinguish 3 regions of human parietal cortex. *Cereb. Cortex* **16**, 1418–1430. (doi:10.1093/cercor/bhj079)

30 Kroliczak, G., Piper, B. J. & Frey, S. H. 2011 Atypical lateralization of language predicts cerebral asymmetries in parietal gesture representations. *Neuropsychologia* **49**, 1698–1702. (doi:10.1016/j.neuropsychologia.2011.02.044)

31 Johnson-Frey, S. H., McCarty, M. & Keen, R. 2004 Reaching beyond spatial perception: effects of intended future actions on visually guided prehension. *Vis. Cogn.* **11**, 371–399. (doi:10.1080/13506280344000329)

32 Elsinger, C. L. & Rosenbaum, D. A. 2003 End posture selection in manual positioning: evidence for feedforward modeling based on a movement choice method. *Exp. Brain Res.* **152**, 499–509. (doi:10.1007/s00221-003-1573-7)

33 Cohen, R. G. & Rosenbaum, D. A. 2004 Where grasps are made reveals how grasps are planned: generation and recall of motor plans. *Exp. Brain Res.* **157**, 486–495. (doi:10.1007/s00221-004-1862-9)

34 Keen, R. 2011 The development of problem solving in young children: a critical cognitive skill. *Annu. Rev. Psychol.* **62**, 1–21. (doi:10.1146/annurev.psych.031809.130730)

35 Chen, Y. P., Keen, R., Rosander, K. & von Hofsten, C. 2010 Movement planning reflects skill level and age changes in toddlers. *Child Dev.* **81**, 1846–1858. (doi:10.1111/j.1467-8624.2010.01514.x)

36 Claxton, L. J., Keen, R. & McCarty, M. E. 2003 Evidence of motor planning in infant reaching behavior. *Psychol. Sci.* **14**, 354–356. (doi:10.1111/1467-9280.24421)

37 Claxton, L. J., McCarty, M. E. & Keen, R. 2009 Self-directed action affects planning in tool-use tasks with toddlers. *Infant Behav. Dev.* **32**, 230–233. (doi:10.1016/j.infbeh.2008.12.004)

38 McCarty, M. E., Clifton, R. K. & Collard, R. R. 1999 Problem solving in infancy: the emergence of an action plan. *Dev. Psychol.* **35**, 1091–1101. (doi:10.1037/0012-1649.35.4.1091)

39 McCarty, M. E. & Keen, R. 2005 Facilitating problem-solving performance among 9- and 12-month-old infants. *J. Cogn. Dev.* **6**, 209–230. (doi:10.1207/s15327647jcd0602_3)

40 Ostry, D. J., Gribble, P. L. & Gracco, V. L. 1996 Coarticulation of jaw movements in speech production: is context sensitivity in speech kinematics centrally planned? *J. Neurosci.* **16**, 1570–1579.

41 Abbs, J. H., Gracco, V. L. & Cole, K. J. 1984 Control of multimovement coordination: sensorimotor mechanisms in speech motor programming. *J. Mot. Behav.* **16**, 195–231.

42 Arbib, M. A. 1981 Perceptual structures and distributed motor control. In *Handbook of neurophysiology: motor control, part 2* (ed. V. B. Brooks), pp. 1449–1480. Bethesda, MD: APA.

43 Rosenbaum, D. A. & Jorgensen, M. J. 1992 Planning macroscopic aspects of manual control. *Hum. Mov. Sci.* **11**, 61–69. (doi:10.1016/0167-9457(92)90050-L)

44 Weiss, D. J., Wark, J. D. & Rosenbaum, D. A. 2007 Monkey see, monkey plan, monkey do: the end-state comfort effect in cotton-top tamarins (*Saguinus oedipus*). *Psychol. Sci.* **18**, 1063–1068. (doi:10.1111/j.1467-9280.2007.02026.x)

45 Chapman, K. M., Weiss, D. J. & Rosenbaum, D. A. 2010 Evolutionary roots of motor planning: the end-state comfort effect in lemurs. *J. Comp. Psychol.* **124**, 229–232. (doi:10.1037/a0018025)

46 Nelson, E., Berthier, N. E., Metevier, C. M. & Novak, M. A. 2011 Evidence for motor planning in monkeys: rhesus macaques select efficient grips when transporting spoons. *Dev. Sci.* **14**, 822–831. (doi:10.1111/j.1467-7687.2010.01030.x)

47 Povinelli, D. J. 2000 *Folk physics for apes: the chimpanzee's theory of how the world works.* Oxford, UK: Oxford University Press.

48 Short, M. W. & Cauraugh, J. H. 1999 Precision hypothesis and the end-state comfort effect. *Acta Psychol.* **100**, 243–252. (doi:10.1016/S0001-6918(98)00020-1)

49 Prescott, T. J., Bryson, J. J. & Seth, A. K. 2007 Introduction. Modelling natural action selection. *Phil. Trans R. Soc. B* **362**, 1521–1529. (doi:10.1098/rstb.2007.2050)

50 Marangon, M., Jacobs, S. & Frey, S. H. 2011 Evidence for context sensitivity of grasp representations in human parietal and premotor cortices. *J. Neurophysiol.* **105**, 2536–2546. (doi:10.1152/jn.00796.2010)

51 Martin, K., Jacobs, S. & Frey, S. H. 2011 Handedness-dependent and -independent cerebral asymmetries in the anterior intraparietal sulcus and ventral premotor cortex during grasp planning. *Neuroimage* **57**, 502–512. (doi:10.1016/j.neuroimage.2011.04.036)

52 Hopkins, W. D. & Pearson, K. 2000 Chimpanzee (*Pan troglodytes*) handedness: variability across multiple measures of hand use. *J. Comp. Psychol.* **114**, 126–135. (doi:10.1037/0735-7036.114.2.126)

53 Hopkins, W. D. & Russell, J. L. 2004 Further evidence of a right hand advantage in motor skill by chimpanzees (*Pan troglodytes*). *Neuropsychologia* **42**, 990–996. (doi:10.1016/j.neuropsychologia.2003.11.017)

54 Annett, M. 2006 The distribution of handedness in chimpanzees: estimating right shift in Hopkins' sample. *Laterality* **11**, 101–109.

55 Coren, S. & Porac, C. 1977 Fifty centuries of right-handedness: the historical record. *Science* **198**, 631–632. (doi:10.1126/science.335510)

56 Steele, J. & Uomini, N. 2005 Humans, tools and handedness. In *Stone knapping: the necessary conditions for a uniquely hominine behaviour* (eds B. Bril & V. Roux), pp. 217–239. Cambridge, UK: MacDonald Institute.

57 Bar, M. 2009 Predictions: a universal principle in the operation of the human brain. Introduction. *Phil. Trans. R. Soc. B* **364**, 1181–1182. (doi:10.1098/rstb.2008.0321)

58 Faisal, A., Stout, D., Apel, J. & Bradley, B. 2010 The manipulative complexity of Lower Paleolithic stone toolmaking. *PLoS ONE* **5**, e13718. (doi:10.1371/journal.pone.0013718)

Research

Functional mastery of percussive technology in nut-cracking and stone-flaking actions: experimental comparison and implications for the evolution of the human brain

Blandine Bril[1,*], **Jeroen Smaers**[2], **James Steele**[2], **Robert Rein**[3],
Tetsushi Nonaka[4], **Gilles Dietrich**[1,5], **Elena Biryukova**[6],
Satoshi Hirata[7] **and Valentine Roux**[8]

[1]*École des Hautes Études en Sciences Sociales—Groupe de recherche 'Apprentissage et Contexte',
190 Avenue de France, 75013 Paris, France*
[2]*AHRC Centre for the Evolution of Cultural Diversity, Institute of Archaeology, University College London,
31-34 Gordon Square, London WC1H 0PY, UK*
[3]*Department of Neurology, Institute of Health Promotion and Clinical Movement Science, German Sport
University Cologne, Am Sportpark Müngersdorf 6, 50933 Köln, Germany*
[4]*Research Institute of Health and Welfare, Kibi International University, 8 Iga-machi, Takahashi,
Okayama, Japan*
[5]*Université Paris Descartes, 1 rue Lacretelle, 75015 Paris, France*
[6]*Institute of Higher Nervous Activity and Neurophysiology, Russian Academy of Sciences,
117485 Moscow, Russia*
[7]*Great Ape Research Institute, Tamano, Okayama 706-0316, Japan*
[8]*CNRS, Maison de l'Archéologie et de l'Ethnologie, Préhistoire et Technologie (UMR 7055),
21 Allée de l'Université, 92023 Nanterre cedex, France*

Various authors have suggested behavioural similarities between tool use in early hominins and chimpanzee nut cracking, where nut cracking might be interpreted as a precursor of more complex stone flaking. In this paper, we bring together and review two separate strands of research on chimpanzee and human tool use and cognitive abilities. Firstly, and in the greatest detail, we review our recent experimental work on behavioural organization and skill acquisition in nut-cracking and stone-knapping tasks, highlighting similarities and differences between the two tasks that may be informative for the interpretation of stone tools in the early archaeological record. Secondly, and more briefly, we outline a model of the comparative neuropsychology of primate tool use and discuss recent descriptive anatomical and statistical analyses of anthropoid primate brain evolution, focusing on cortico-cerebellar systems. By juxtaposing these two strands of research, we are able to identify unsolved problems that can usefully be addressed by future research in each of these two research areas.

Keywords: hominin; chimpanzee; Oldowan; nut cracking; experimental archaeology;
cortico-cerebellar

1. INTRODUCTION

Archaeological evidence suggests that tool use has been fundamental to hominin life for at least 2.6 Myr [1,2] and probably more [3,4]. Stone knapping represents the earliest known instance of toolmaking and tool use by early hominins [5–10]. Stone tool production has therefore become diagnostic of the cognitive ability and motor skills of extinct hominins [11–14]. Following the first scientific report of chimpanzee tool use by Jane Goodall in 1964 (the use of stripped leaf stalks for termiting, sticks for ant-dipping and leaves for drinking and self-wiping [15]), numerous observations have also been made of tool use in non-human primates in the wild, as well as in controlled experimental conditions in captivity. Use of tools to crack nuts in forest-dwelling chimpanzee groups has now been widely attested (early reports included Beatty [16] in Liberia; Struhsaker & Hunkeler [17], Rahm [18], Boesch [19] in the Tai

forest, Ivory Coast; and Sugiyama & Koman [20] and Sugiyama [21] in Guinea).

In contrast with the use of stone hammers to pound and crack the casings of hard food objects, stone flaking to produce cutting tools appears to be a uniquely hominin cultural trait [12]. Toth & Schick [22] suggest that modern wild chimpanzees have not acquired stone-flaking traditions because none of their feeding objects need to be accessed by cutting. In contrast, the hominins associated with Oldowan stone tools were regularly feeding on animal prey for which such cutting tools and techniques were essential [23]. The earliest direct archaeological evidence of food item processing using Oldowan stone tools is from animal bones, which show cut marks associated with stripping off edible soft tissue and also fractures associated with cracking them open to obtain edible bone marrow. Using stone tools to cut animal soft tissue is attested from marks on the surfaces of bones in the earliest archaeological record (2.5–2.6 Myr BP, associated with the extinct hominin *Australopithecus garhi* [24]; possibly also at 3.4 Myr BP and associated with *Australopithecus afarensis* [3], but see [25]).

Various authors have suggested behavioural similarities between tool use in early hominins and chimpanzee nut cracking, where nut cracking might be interpreted as a precursor of more complex stone flaking [20,26–28]. But does the production of stone cutting tools require different skills, and different levels of functional understanding, than the use of stone hammers to fracture casings of hard food objects? If so, then Oldowan stone tool production may be predictive of significant differences in the associated cognitive abilities and motor skills of modern chimpanzees and extinct hominins.

In this paper, we bring together and review two separate strands of research on chimpanzee and human tool use and cognitive abilities. Firstly, and in the greatest detail, we review our recent experimental work on behavioural organization and skill acquisition in nut-cracking and stone-knapping tasks, highlighting similarities and differences between the two tasks that may be informative for the interpretation of stone tools in the early archaeological record. Secondly, we outline a model of the comparative neuropsychology of primate tool use and review recent descriptive anatomical and statistical analyses of anthropoid primate brain evolution, focusing on the chimpanzee–human comparison. By juxtaposing these two strands of research, we are able to identify unsolved problems that can usefully be addressed by future research in each of these two research areas.

2. FUNCTIONAL PARAMETERS OF PERCUSSIVE TECHNOLOGIES

Percussion can be loosely defined as 'a forceful, muscle-driven striking of one body against another' [28, p. 342], but this definition does not specify the way in which force is controlled to transform an object. A tool-assisted percussive task involves delivering a blow or a series of blows with an object, typically held in the hand, in such a way that all the parameters and constraints of the task are met. This definition may be applied to activities such as hammering a

nail, drumming, hitting a golf ball, cracking a nut, flaking a stone, etc. Mechanically speaking, success depends on the properties of the object being struck and on the value of the momentum of the tool (hammer, drumstick, golf club, etc.), which is defined as the product of its mass (m) and its velocity (v). For a biological system, the efficiency of a blow can be defined in terms of potential and kinetic energy. An object held in a person's hand has potential energy—energy of position—which converts to kinetic energy—the energy of motion—if the actor lets it fall to the ground. If no additional—i.e. muscular—energy is added to the system, the sum of kinetic and potential energy stays constant. For a biological system, an energy-efficient blow is one in which the minimum of muscular energy is added to the system to achieve the task goals. Indeed, for typical learned movements, the external forces and passive forces of reaction in the joints are by far the most used in movement construction. Consequently, a minimum of muscular energy is added to the system to achieve the task goal.

To characterize the skills needed for percussive techniques of nut cracking and stone flaking, we differentiate four layers of parameters [29] defining (respectively) the functional and deterministic task constraints, the parameters under the control of the actor performing the task and the parameters that determine effective regulation of these control parameters and movement parameters (figure 1). The *functional parameters* specify the topology of the task through relevant geometric and dynamical parameters, including kinetic energy, point of percussion, angle of blow and (for stone knapping) exterior platform angle. With regard to the dynamical parameter of kinetic energy ($E_k = 1/2\ mv^2$), the layer of *control parameters* specifies two parameters, the velocity (v) at impact and the mass of the hammer (m) that includes its substance and density, which are typically under the control of the actor. Finally, given a specific hammer, velocity can be regulated through various strategies, which depend on the actor. For example, the movement may be of large amplitude, relying on high potential energy and low muscular energy, or the opposite, with a small amplitude but with a large additional input of muscular energy. *Regulatory parameters* can therefore vary between actors who use alternative bodily movements to achieve the same functional output [30]. Movement parameters are those parameters that can be recorded and that allow the computation of regulatory and control parameters. These include kinematics, kinetics and muscular parameters that can be recorded through various technical means. This level of analysis will not be discussed here. For more discussion, see Biryukova & Bril [30].

(a) *Nut-cracking techniques*
In the case of *nut cracking*, the blow must be delivered in such a way that the shell cracks leaving the kernel intact. To achieve this goal, the right amount of kinetic energy must be generated and transferred to the nut in order to produce an adequate deformation of the shell so that it breaks. This depends on the hardness of the nut shell: if the kinetic energy is too high, the nut will be smashed and the kernel may be ruined, while if the
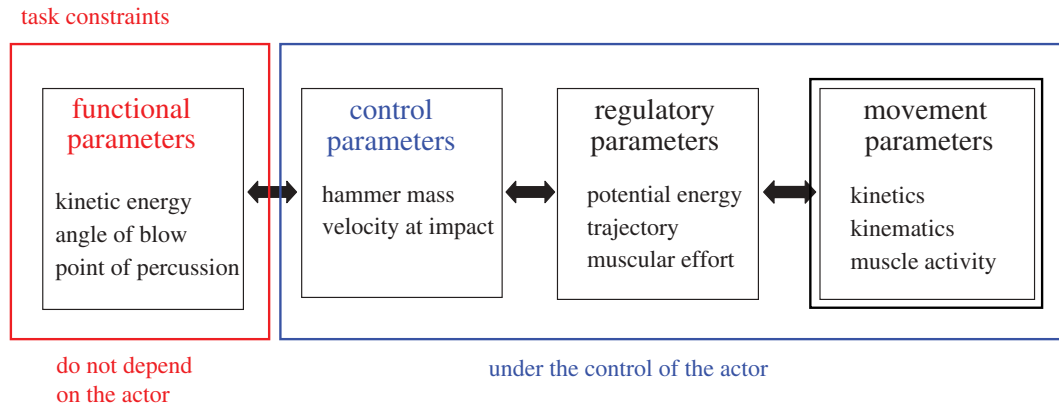
Figure 1. The three layers proposed for percussive tasks. Except for the exterior platform angle, all the parameters in some way or another have to be controlled in any percussive task. Only movement parameters are recorded and allow for computation of regulatory and control parameters (adapted from Bril *et al.* [16]).

kinetic energy is too low, the shell will not crack. Nut shells have evolved to be resistant to fracture, to protect the seed from predation. Consequently, their strength is largely independent of the point on the external surface receiving the percussive strike, which must therefore simply apply enough force to induce fracturing. Nut shells can be very hard indeed—macadamia nuts, for example, have an elastic modulus (a measure of the material's stiffness, or resistance to permanent deformation under compressive loading) of the order of $2-6 \, \text{kN} \, \text{mm}^{-2}$ [31], and require a force of the order of 2 kN to fracture them [32–34], while typical orally processed primate foodstuffs given in captivity have an elastic modulus in the range $0.1-350 \, \text{N} \, \text{mm}^{-2}$ [35]. The force required to fracture nuts of the species used by wild chimpanzees varies with the nut species and condition: typical forces required to fracture such nuts range from 2.8 kN for *Coula edulis* to 8.1 kN for *Parinari excelsa*, and between 9.7 and 12.5 kN for *Panda oleosa* ([36]; cf. [37]). Koya [38], however, found that with a hammer and anvil technique, repeated blows of much less than the force required individually to induce fracturing will still cause the palm nut shell to fail, because of the induction of micro-fractures and subsequent fatigue failure. This suggests that repetitive pounding, rather than attempting to fracture a nut by a single forceful strike, may be energetically a less costly strategy as well as one less likely to accidentally crush the kernel.

Among non-human primates, banging the food object against a hard surface is a frequently observed technique to crack open nuts. Controlled experiments as well as observations in the wild demonstrate the very fine adjustment to the constraints of the task by capuchin monkeys (e.g. [39–41]) as well as chimpanzees, to enable them to reach their goal. Wild chimpanzees (e.g. [42]) and contemporary human foragers are both reported as using a hammer and anvil technique, in which the nutshell is forcibly compressed between two hard surfaces. For energetic efficiency with the hammer and anvil technique, the blow must be elastic, the total impulse being constant before and after the blow so that, in theory, all forces are used to generate the deformation of the shell in such a way that it cracks. If the blow is a non-elastic one, a part or all of the energy will be dissipated, and it will be difficult

to crack open the nut. For example, if the nut was lying on a soft surface or anvil, the energy would be absorbed by the support and the nut would not crack or would need a very high amount of energy in order to reach its breaking point. In addition, the direction of the blow must be more or less perpendicular to the surface on which the nut rests, since otherwise it would be displaced laterally (the energy being used to increase the velocity of the nut, and not to attain the goal of cracking it!). Fracturing these nuts using a stone hammer and anvil requires however only the stable placement of the nut on a sufficiently hard anvil, and delivery of a blow with a velocity vector approximately perpendicular to the plane of the anvil, of sufficient force to compress the nut and induce fracturing.

The early archaeological record also contains evidence of stone tool use in pounding or cracking open hard food objects, using techniques that are much more closely analogous to chimpanzee tool use than the production of Oldowan stone flakes. Mora & de la Torre [43] have reanalysed the stone tools from the lowest levels at Olduvai Gorge and suggested that the majority may relate to pounding hard food objects (bones, nuts), and not to producing stone flakes. Stone tool use to crack open bones would have been an essential technique in the Oldowan hominin repertoire. Pitted stones that could have been used for bone and nut cracking or for bipolar stone flaking are found in Oldowan levels at Olduvai Gorge (Tanzania) and also at Melka Kunture (Ethiopia), although the nut-cracking function is not yet directly attested for sites older than Gesher Benot Ya'aqov (Israel), where nut remains have been found in association with stones with surface pitting, and which dates to oxygen isotope stage 19 (*ca* 780 000 BP; [44]).

Nut cracking by a stone hammer and anvil is also found in contemporary human hunter–gatherers. The !Kung of the Kalahari desert, one of the quintessential hunter–gatherer groups of modern ethnography, process mongongo nuts for about 28 per cent of their food intake [45]. These nutshells are extremely hard and must be processed by cracking them open in this way, sometimes preceded by roasting to make the shells more brittle. Stone tools used for cracking mongongo nuts make up a disproportionately large fraction

of the tools found discarded at !Kung campsites [46]. The skills involved are learned and continue to improve with experience well into adulthood. Bock [47], in a cross-sectional study conducted in 1994, experimentally measured the rate at which !Kung foragers could extract intact mongongo nut kernels from their outer shells using stone tools, and found significant effects of age (but not independently of strength) on processing rates, with efficiency in mongongo nut cracking continuing to improve through the teenage years and twenties and peaking among adults aged in the thirties and forties. Similar techniques continue to be used by subsistence and small-scale farmers: Koya & Faborode [48] found that forces of about 2 kN or higher were sufficient to induce fractures in palm nuts that are used to extract palm oil (*Elaeis guineensis*), and which are characteristically fractured individually by Nigerian peasant farmers who 'break the nuts, one at a time, between two stones judging the magnitude of the applied force by experience' [48, p. 471]. However, very often it is observed that cracking open a nut necessitates several strikes.

(b) Stone-flaking techniques

The fracture mechanics of the stone-flaking task are very different. Fine-grained stone typically has greater compressive than tensile strength (i.e. it is brittle [49]), which means that despite its hardness it can be fractured easily if force is applied in the right location and direction. The two main modes of fracture initiation to be considered here are wedge-fracturing and conchoidal fracturing [50,51].

Wedge fractures are initiated when force is applied and either detrital particles become wedged into a pre-existing flaw on the core surface, or the core surface is plastically deformed by forceful contact from a hard and sharp indenter; in both cases, the wedging causes crack initiation. This is the predominant mode of stone fracturing when the force is applied at a location far from a platform edge, or if the edge angle exceeds 90°, or if the core has many internal flaws [50, p. 688]. It is the typical fracture mode in bipolar flaking, where a pebble is placed on a hard anvil and hit with a hard hammer stone until it splits ('the method of bipolar flaking is much like cracking a nut with a hammer' [52, p. 131]).

Wedge fracturing corresponds to what Pelegrin [53] calls 'split breaking', and is the mode of flaking that seems to characterize the solutions that the captive bonobo Kanzi has developed when taught stone knapping. In split breaking, if a sufficient load is applied, the stone will break no matter how and where it is applied: essentially, all that is required is the localized application of sufficient force to the core to initiate a wedge. Consequently, the properties of the flake to be detached cannot be finely controlled; this solution is therefore fundamentally no more difficult than the solution of the nut-cracking task. When Kanzi had to face a situation where the goal was to cut through a cord to open a box containing a desirable food [54,55], he succeeded in discovering a way to produce a chip with a sharp edge. However, while Kanzi was encouraged to produce flakes with a sharp edge

through direct percussion using a hard stone hammer to strike the core, he developed his own technique to get a sharp-edged piece of stone that would perfectly fit his goal by throwing the core against a hard surface. Although he was trained for quite a few years, he has never developed a technique that allows him to produce conchoidal flakes intentionally. Throwing does not seem to be a common technique to open nuts or fruits although theoretically it should have comparable functional properties to other percussive techniques, and could allow for the production of greater kinetic energy at contact (the full lever action of the whole arm could be brought to bear, with a reduced requirement for accuracy in the trajectory of the throw). Marchant & McGrew [28] relate the case of chimpanzees in the Parc National du Niokolo-Koba in Senegal throwing baobab fruits against an anvil to smash them open. However, this technique may be less likely to work on fruit and nut shells than on stone cores, because these materials differ in their elastic properties (likelihood of bouncing off a hard surface).

Controlled conchoidal fracturing, in contrast, requires a fuller understanding of the behaviour of the stone core under dynamic impact loading ([50]; see [56]). Hence, successful flake detachment by conchoidal fracture requires finding the appropriate point of percussion and achieving a sufficiently high degree of striking precision. Therefore, the properties of the flake to be detached can be strictly controlled by the knapper [56]. In terms of the fracture mechanics, conchoidal fracture is typically initiated with a partial Hertzian cone, caused by compression of the core at the point of impact by a hard indenter, followed by a crack propagation phase (figure 2). According to Cotterell & Kamminga [50,51], the intrinsic stiffness of the raw material means that cracks tend to propagate parallel to the plane of the external flake surface, which means that detachment is somewhat insensitive to the precise angle of contact between the platform and the percussor; when the direction and magnitude of applied force cause the crack to propagate away from that plane, the intrinsic stiffness of the material can still bring it back into that plane, producing characteristic ripples or undulations on a flake's ventral surface. In Cotterell and Kamminga's experiments, flake length (dimensionless, scaled to platform depth at the point of percussion) was found to have a greater dependence on the external platform angle and on the morphology of the dorsal surface of the developing flake (however, new experiments by Dibble & Rezek [57] suggest that the angle of the blow can also be a relevant factor). Unsuccessful flake detachment can occur when the force applied is insufficient for the given location on the core. An excess of outwardly directed force in the percussive strike may cause a hinge fracture to develop [50]. Application of insufficient force may cause the flaking energy to be consumed before the crack has propagated to the distal edge of the core, causing a step fracture; this may occur for example at locations on the core with a wide flat external surface. Hinge and step fractures are characteristic of novice stone knappers (e.g. [58]), and reflect poor understanding of the appropriate force needed to detach a flake with a
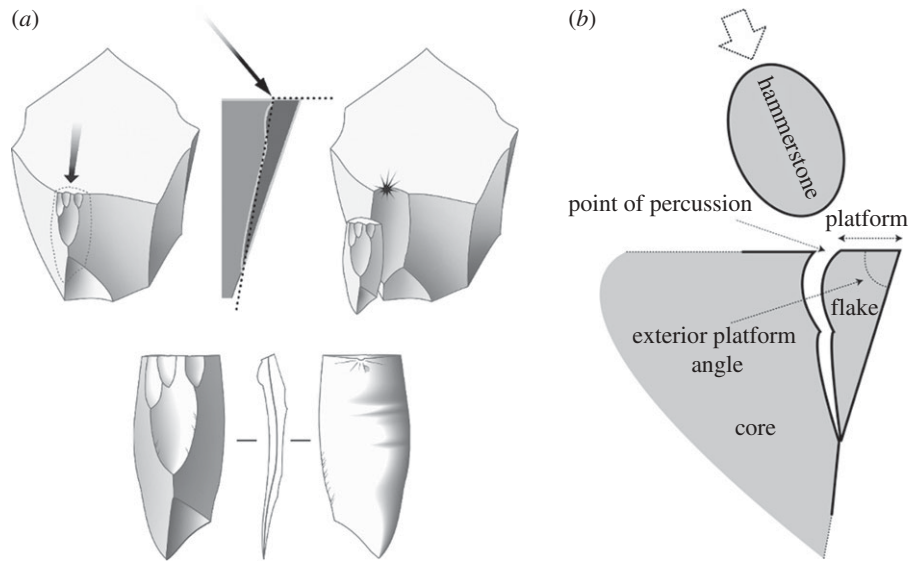
Figure 2. (*a*) Conchoidal fracture resulting from an angle of percussion near 40–50°, and an exterior edge angle at 70–80°. Bottom showing the characteristic feature of a flake: the swell at the point of contact or bulb of percussion is clearly visible (reproduced with permission from Pelegrin [53, p. 24]). (*b*) Flaking terminology.

feathered termination from a particular location on the core's external surface. Experienced knappers will typically choose the location much more carefully at which to apply appropriate force, and may exploit these properties of the raw material: for instance, in blade removal, a longitudinal ridge on the dorsal surface of the intended blade will be exploited because it prevents the crack spreading laterally during stiffness-controlled crack propagation [50].

In the case of *percussive conchoidal flaking*, similar to nut cracking, the blow must be an elastic blow, delivered in such a way that a flake is detached from the core responding to the mechanism of the conchoidal fracture. However, the constraints of the task, i.e. the functional parameters described above, are more numerous than in nut cracking. The shape and size of the flake depend on several parameters: the exterior platform angle, the point of percussion, the angle of the blow relative to the platform and the kinetic energy that initiates the fracture. A peculiarity of the kinetic energy necessary to produce a conchoidal fracture is the existence of a threshold value. Once a minimum effective quantity of kinetic energy is produced, an increase in this value has no impact on the flake produced, except that a value far too large may cause the flake to fragment into many pieces (reference in [29]). As such, the characteristics of the flake (its dimensions and form) depend on the convergence of multiple interrelated variables [56,57,59].

A variety of techniques for conchoidal flaking are known ethnographically, and studied experimentally (e.g. [60, p. 31]), including direct percussion with a hard or soft hammer; indirect percussion with a punch either interposed between the hard hammer and the core, or located on the opposite side of the core to the hammer ('counter-blow'); as well as non-percussive pressure flaking techniques. In this paper, we are mainly concerned with the techniques of direct percussion with a hard hammer, which were the primary flaking techniques used in Oldowan toolmaking.

3. MASTERING THE FUNCTIONAL PARAMETERS OF PERCUSSIVE ACTIONS: EXPERIMENTAL EVIDENCE

We now summarize a series of experiments designed to establish some of the parameters involved in skilled action in percussive tool-using tasks.

(a) *Material and methods*
We briefly present our methods and subject populations here; more details can be found in our papers reporting the primary results [29,32,33,56,61–66]. We designed experiments to investigate how actors of various levels of expertise develop specific behavioural traits concerning movement precision, flexibility, smoothness, regularity and optimization [67,68] (cf. figure 3). In the stone-knapping experiments in Khambhat (Gujarat, India), craftsmen of different levels of expertise have been asked to knap beads of different shapes and raw material, using an indirect percussion technique and with hammers having different properties. In the Oldowan replication study, modern experimental knappers having various amounts of practice were asked to produce conchoidal flakes of different sizes with hammers of various weights. The nut-cracking experiments with humans as well as chimpanzees were based on the same rationale, cracking nuts of different hardness with hammers of various weights. Children from 5 to 12 years of age and adults stood for actors of different levels of expertise.

To be able to compare the movement in human stone knapping and in nut cracking by humans as well as chimpanzees, we used recording techniques that could provide data as similar as possible in all cases. Humans' knapping and nut-cracking movements were recorded with an electromagnetic recording system (either a Polhemus system or a Flock of Bird system; Ascension Technology Corporation) and an accelerometer for the Indian craftsmen. Chimpanzees' nut-cracking movements were recorded using two cameras positioned on the right and left of the animal with an angle of approximately 100°. Figure 3 shows examples of these experimental settings.
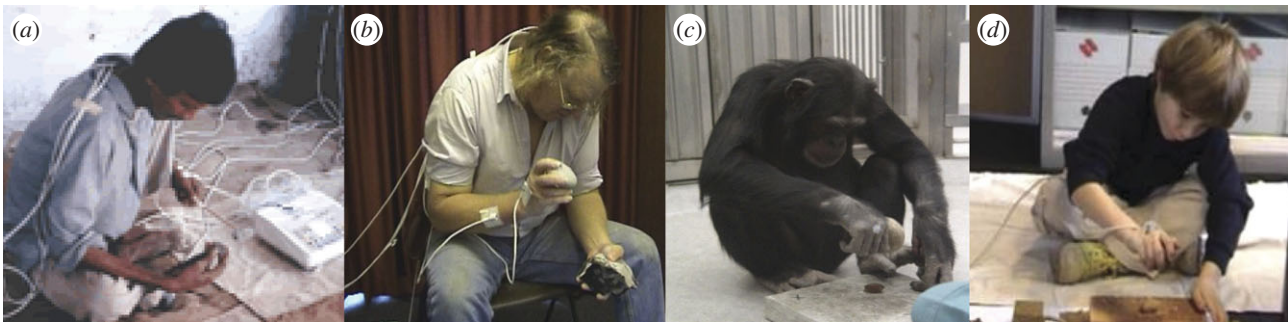
Figure 3. Illustration of the experimental setting for the different experiments. (*a*) Craftsman knapping a bead by means of indirect percussion by counter blow. (*b*) Modern experimental knapper producing a flake by direct hard-hammer percussion. (*c*) Chimpanzee cracking a nut (GARI, Japan Copyright © S. Hirata). (*d*) An 8-year-old child cracking a nut.
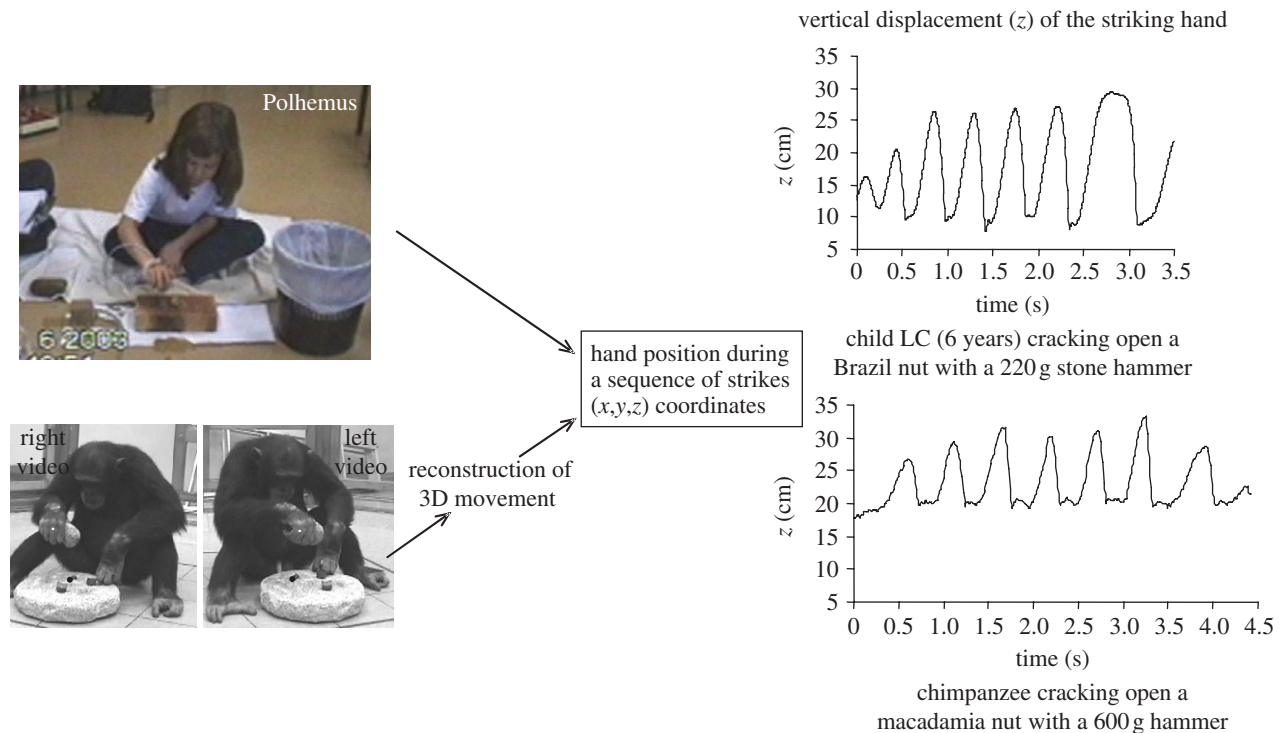
Figure 4. Recording and reconstruction of the hand movement for humans from the electromagnetic sensors, and from video cameras for the chimpanzees.

For both experiments on stone knapping with human adult subjects, sensors were attached with tape to the acromion, the exterior surface of the humerus, the posterior surface of the lower arm and to the dorsal surface of the hand following the procedure used by Biryukova *et al.* [69]. For children nut cracking, one sensor only was used (on the dorsal surface of the hand). As active or passive markers cannot be used for chimpanzees, we used a two-camera video-based system that allows reconstructing three-dimensional movement (figure 4). The reconstructed three-dimensional movement of the striking hand and the computation of functional parameters of the striking action necessitated frame-by-frame analysis [32,33].

(b) Behavioural results I: control of kinetic energy in adaptation to task conditions

Kinetic energy, which is a key functional parameter to percussive tasks, involves two control parameters that

are under the control of the actor: the mass of the hammer, and the velocity of the hammer at the time of contact. In the wild, chimpanzees have been observed selecting the appropriate tool, i.e. hammers and anvils of particular size, shape and materials, suggesting that they apprehend the functional properties of the nut-cracking task [20,21,42,70]. This capacity of selecting functional tools has also been shown in capuchins [40,41,71]. However, this capacity to perceive the affordances of objects as potential tools needs to be learned through experience. In an experiment where children were offered a set of 21 objects as potential tools of various degrees of functionality [61], potentially functional hammers represented 52 per cent of the objects chosen by 3 year olds, 90 per cent of the choices of 4/5 and 6/7 year olds and almost 100 per cent of choices of older children. While we have not done similar experiments exploring stone tool choice for stone knapping, experiments with Indian craftsmen also suggest that the recognition of subtle contrasts in suitability of possible hammers depends on the level of expertise [72].
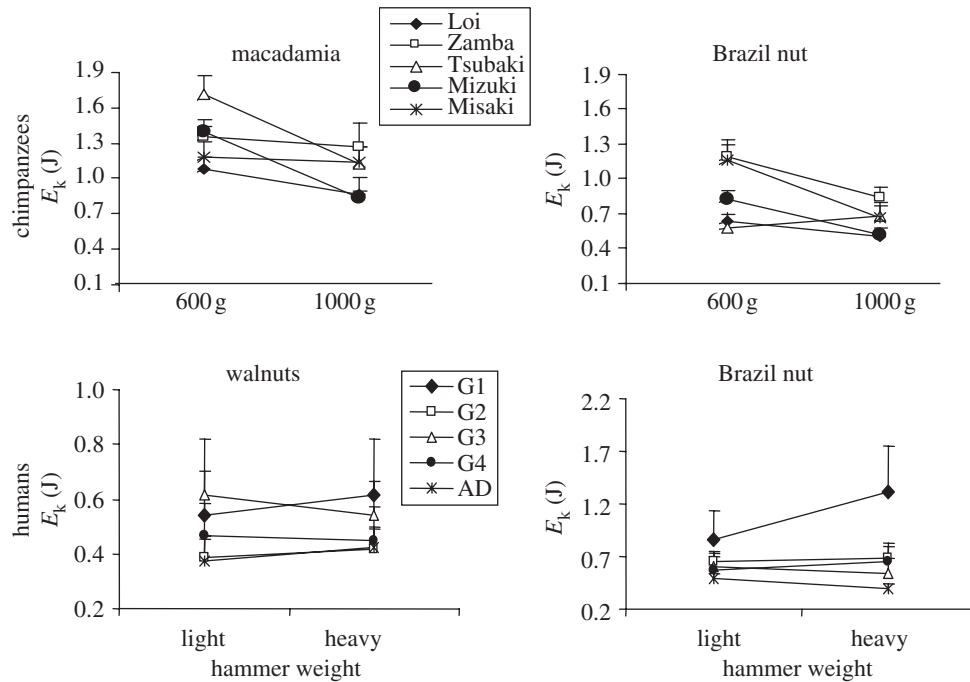
Figure 5. Kinetic energy values at the time of contact of the hammer with the nut for chimpanzees and humans when they crack nuts of two different species (different hardness) with hammers of different weights. For the chimpanzees, values are given for each chimpanzee. For humans, data are averaged for each age group, children from 5–6 years (G1) to 11–12 years (G4) and adults (AD; adapted from Bril *et al.* [32] and Foucart [64]).

Once the hammer has been chosen, the kinetic energy depends only on the velocity at impact, which is under the control of the actor. To what degree are chimpanzees and humans able to adapt to the hammer properties to crack nuts or to knap stone? We conducted experiments to evaluate the capacity of the actor to cope with the constraints of the task by comparing the values of kinetic energy produced when the actor had to use hammers of different weights [29,32,64]. All actors (humans and chimpanzees) were able to modify the velocity they produce to the changing constraints of the task; in all cases—i.e. chimpanzees or children cracking nuts and adult humans flaking stone—the velocity was systematically greater when using a lighter hammer. However, in both tasks and in both species, a higher level of expertise was associated with a reduced velocity at impact. In the nut-cracking task, in both species, there was also evidence of an ability to adjust velocity to differences in hammer mass in order to deliver a constant level of kinetic energy at impact (figure 5; [32,64]). When looking at stone flaking, the results are quite different. While all knappers showed greater velocity when using a lighter hammer, these variations do not end up in the production of the same kinetic energy when using a light or a heavy hammer. The adaptation to the hammer weight that was observed for velocity does not transfer to kinetic energy, except for experts (figure 6; [29,56]).

Our results undoubtedly show that experience is a key criterion in the understanding of the constraints of the task. Experts display exactly the same kinetic energy while both intermediates and novices displayed greater kinetic energy with lighter hammers, a result that may be compared with those obtained with chimpanzees. In another set of experiments, novices produced values of kinetic energy more than three
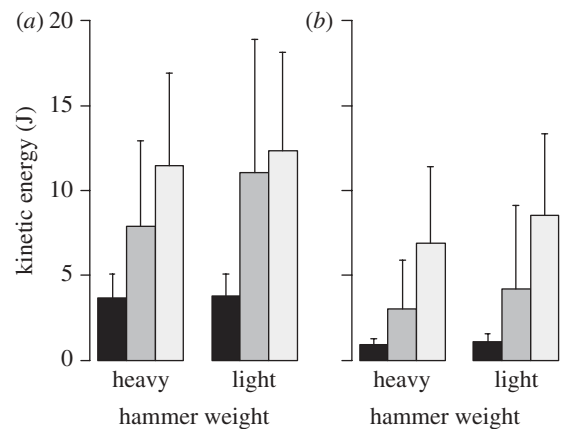
Figure 6. Kinetic energy at point of contact between the hammer and the core, when the task was to knap a large flake and the subjects had been previously classified as experts, intermediates and novices (adapted from Bril *et al.* [16]). Black grey bars, expert; light grey bars, intermediate; white bars, novice. (*a*) Large flake; (*b*) small flake.

times greater compared with experts, and produced smaller flakes (figure 7). The dramatic differences in the values of the kinetic energy at impact observed between experts and novices in stone knapping contrast with the relatively small differences observed in adults and children when cracking nuts. Except for the youngest children aged 5–6 years in the more difficult situation (cracking open Brazil nuts), all the participants were able to adjust the kinetic energy delivered to the nut in different conditions of hardness of the nut and weight of the stone hammer. In the same way, in most cases chimpanzees are able to adapt to both hardness of the nut and weight of the hammer in the nut-cracking task, even though in a less fine-tuned manner.
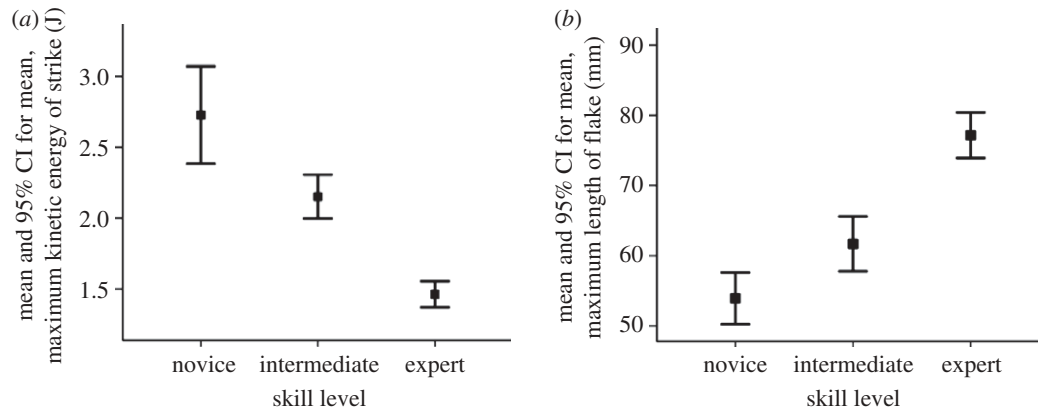
Figure 7. (*a*) Mean values for maximum kinetic energy of strike, and (*b*) for maximum flake length, by skill level. The task was to remove a flake (size unspecified) from a prepared single-platform core. Subjects were two experts making a total of 77 flake removals, three intermediates making a total of 131 flake removals and four novices making a total of 149 flake removals. Greater expertise is associated with less kinetic energy and larger flakes. (From F. Wenban-Smith, B. Bril, G. Dietrich, R. Rein, T. Nonaka & J. Steele 2010, unpublished data.)

We may hypothesize that the number of variables to be controlled in the case of nut cracking is much smaller than in stone knapping. In nut cracking, the amount of kinetic energy must be controlled to reach the breaking point, while stone knapping is characterized by a threshold value that must be discovered by the actor [29,56]. In nut cracking, two functional parameters only have to be controlled, kinetic energy and direction of the blow. As the nut is positioned on an anvil, the direction of the blow is approximately vertical. The exploration process necessary to find out the efficient amount of kinetic energy may be conceived as a succession of approximations that progressively converge to the right value and necessitates the control of one parameter only. The case of knapping conchoidal flakes is entirely different. The interrelationship of numerous variables to succeed in detaching the planned flake makes the task incomparably more difficult, and consequently, the exploration process of the task space will also be tremendously more complex.

(c) Behavioural results II: bimanual coordination

We have also investigated asymmetric bimanual coordination in a stone-knapping task. Previous studies have provided substantial evidence of functional asymmetries between the two hands and their underlying neural structures [73–75]. Yet, the problem of how such asymmetric bimanual activities are organized into the collective behaviour of a bimanual system still remains incompletely understood [66]. In one of the few papers addressing the issue of the coordination between the asymmetric elements, Guiard [76] proposed a 'kinematic chain model' to explain functional coordination in human skilled bimanual actions [75,77]. The essence of Guiard's [76,78] conceptualization is that he considers the two hands as serially assembled, instead of following a parallel assembly pattern. What is implied in serial assembly is that two different layers of activities in two hands are coupled with each other to contribute to the same output. Guiard [76] hypothesized that

the outstanding manipulative efficiency of humans results not only from role differentiation between the two hands or the emergence of handedness but also from the fact that between-hand division of labour is typically functionally nested, with two hands working at two different levels of resolution in a coordinated fashion to yield a common functional outcome. In his model, one hand and/or arm performs movements that Guiard qualifies as high frequency, being more temporally and spatially precise (i.e. being faster and having a narrower target), whereas the other upper limb is low frequency, acting as a stabilizer or support, maintaining the spatial or temporal structure, and moving earlier to define the spatial reference frame. To define the group-level handedness that is specific to humans, Guiard suggested that most humans tend to learn the low-frequency role with the left hand and the high-frequency component with the right hand. Such human population-level right handedness is generally explained by reference to a left hemisphere advantage for fine temporal resolution of sensory input and motor output. Carson [79, p. 481] discusses two possible explanations for this advantage. One is that the left hemisphere may be more efficient in error correction using sensory feedback. The other is that the left hemisphere may permit more precise control of net forces and force durations (compare also [80,81]).

Non-human primates must also be observed using their hands in complex asymmetric bimanually coordinated tasks if the objective is to record hand preferences (e.g. [82]). The task that is most frequently used at present to elicit such behaviours in captive populations is the tube task [83], an extractive feeding task involving an opaque polyvinylchloride (PVC) tube containing smears of peanut butter that can be extracted if one hand holds the tube while part of the other hand is inserted into it. Hopkins *et al.* [83] have found a population-level right-hand preference in the tube task in three separate captive chimpanzee populations, all with large sample sizes, although the ratio of right- to left-handed individuals is lower than in humans—typically 2 : 1 in chimpanzees, as compared with 8 or 9 : 1 in humans—and, furthermore,

there are very much higher frequencies of ambiprefer-ent chimpanzees than of ambipreferent humans [84]. A similar pattern has not been reported in wild chim-panzees in nut-cracking tasks, although this may be because that task is less complex in terms of bimanual coordination.

We have not yet analysed bimanual coordination in chimpanzees in the nut-cracking task. We have how-ever analysed bimanual coordination in human stone knapping, and this provides some insight into the cor-relates of expertise. In this task, the two hands have clearly differentiated functions. The hammering hand needs to be controlled in such a way as to transmit the appropriate amount of kinetic energy at impact with considerable accuracy at the point of percussion. On the other hand, the postural hand has to rotate and adjust the position of a core or rough-out to prepare for the following hammer strike, and stabilize the core or rough-out against the shock of the blow.

Professional craftsmen from two classes of workshops in Khambhat (Gujarat, India) participated in our exper-iment [66]. In addition to carnelian stones typically used in bead production, a new raw material—glass—was also included. Among the sub-goals that make up the task, 30 s sequences of calibration (standardization of crests to prepare for fluting) and fluting (detachment of long crests) were extracted from each trial. Depen-dent variables calculated for each 30 s sequence were used for the statistical analyses. The two sub-goal sequences chosen for analysis have different functional requirements. Fluting, through one strike, determines the overall shape of the product by detaching a long flake. On the other hand, calibration is more of a process of standardizing the crest to prepare for fluting. Each fluting sequence consists of small preparatory move-ments and several forceful strikes to detach long flakes, whereas each calibration sequence consists of a series of detachments of a number of tiny flakes.

Previous studies from Khambhat have shown that the end products produced by high-level expert craftsmen (trained with a longer apprenticeship period) had signi-ficantly greater sphericity and a smoother surface than those produced by low-level expert craftsmen (trained with a shorter apprenticeship period), and that such a group difference was amplified in the novel situation using glass rough-outs [62,63,65]. We studied asym-metric bimanual coordination of professional bead craftsmen from these two skill level groups in a naturalistic situation using recurrence methods [66]. Our key findings are that the movements of the two hands of craftsmen were controlled, reflecting the functional requirements of the task and the roles assumed by each hand, and that the skill level difference appeared in the way they were organized into a unified act. Regarding the func-tional specificity, among others, evidence was found in both groups that the dynamics of the displacement of the hammering hand and that of the postural hand were both relatively stable when glass was used, and that the dynamics of the displacement of the postural hand was relatively stable during fluting compared with calibration. However, only the bimanual movement coordination of highly skilled experts differentiated the functional requirements of different sub-goals. Furthermore, the dynamics of bimanual movement of high-level experts

exhibited more deterministic coupling than that of low-level expert craftsmen. These results suggest that what is acquired in skilled bimanual action is adaptable and flexible nesting of differentiated functions, in which movements of two hands are modulated in such a way as to meet various functional demands of the situation.

(d) *Behavioural results III: understanding of fracture mechanics*

Anticipation is often considered to necessitate a high level of cognitive ability. If we define anticipatory behav-iour as behaviour that prepares for the forthcoming goal, humans as well as other primates must be capable of anticipatory behaviour. The choice of a hammer adapted to the hardness of the nut prior to the actual nut-cracking activity has been observed in capuchins as well as chimpanzees [21,40,42,71]. Anticipation may be observed at the level of the striking action as well. We have been able to show how chimpanzees can modulate the kinetic energy in a sequence of strikes to crack open a macadamia nut [32]. Figure 8 shows the 11 strikes given to a nut before taking the kernel out. While the value of potential energy is constant, the kin-etic energy increases up to a certain value, and remains constant for the last two hits, when probably the shell is broken; then a few low-kinetic-value strikes are given probably to take the bits of shell off. We hypothesize that this striking strategy suggests that the chimpanzee has some 'understanding' of the existence of a breaking point that should not be passed over.

Anticipatory procedures are a great deal more complex when looking at conchoidal flaking. When planning to knap a flake with defined characteristics, a large number of interrelated features of task constraints have to be taken into account, and behaviour has to be adjusted accordingly. In a recent study, Nonaka *et al.* [56] have shown that only very high-level expert knappers are able to produce the flake they intended. The intentions of the knappers were analysed prior to the actual flaking, in terms of the expected shape of the detached flake and the intended percussion point. Results showed that to predict accurately the consequences of a strike entails an acute exploration of the properties of the core and of the hammer stone, to set up an interrelationship among the variables in such a way as to comply with the task con-straints. In our study, only experts with approximately 20 years of part-time experience in replicating archaeo-logical stone tools were able to predict flake dimensions that significantly correlate to the detached flakes. The fact that experts are able to accurately predict the flake to be detached suggests that under such a part-time training and practising regime, years of experience may be necess-ary to understand the constraints of the conchoidal fracture, and that it requires similar amounts of experi-ence to be able to discriminate the feature of a core and the functional properties of the action that affect the morphology of the flake (figure 9).

4. BRAIN EVOLUTION IN HUMANS AND CHIMPANZEES: ISSUES RELEVANT TO TOOL USE

As outlined above, we have identified similarities and contrasts in the structure and cognitive demands of
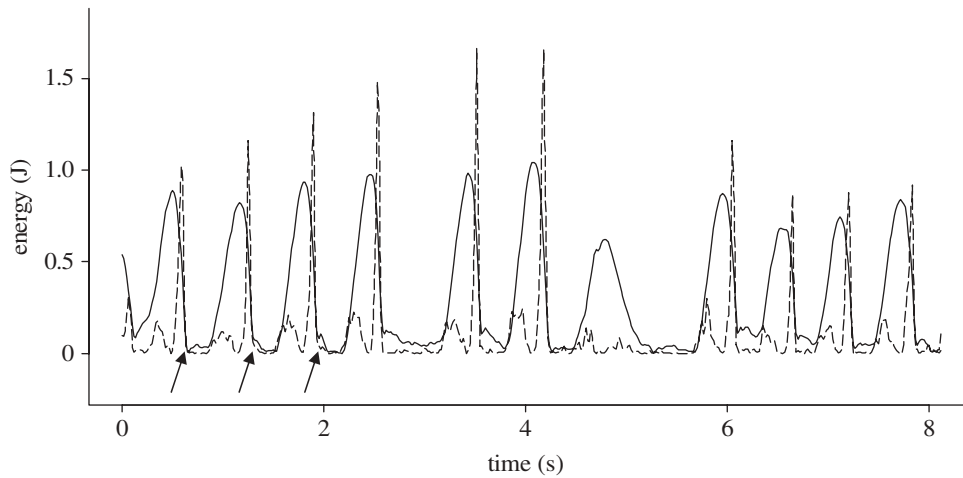
Figure 8. Potential (E_p, solid lines) and kinetic energy (E_k, dashed lines) during a sequence of strokes by a chimpanzee cracking a macadamia nut with a 300 g hammer. The arrows mark the time of impact of the stone hammer on the nut (adapted from Bril *et al.* [32]).
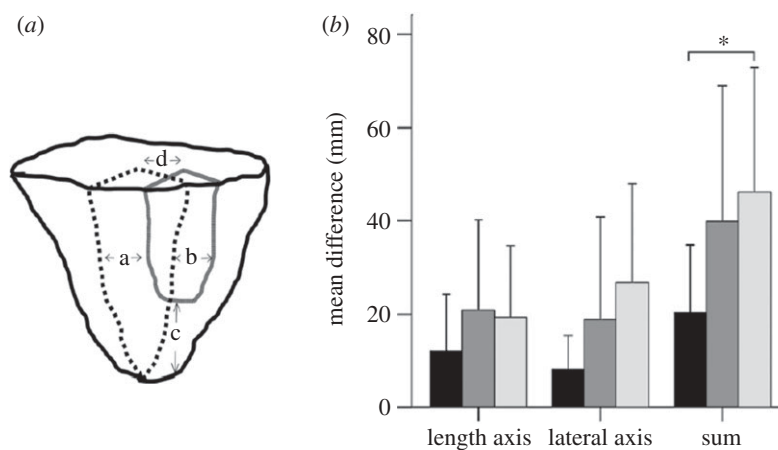
Figure 9. (*a*) The measurements made on predicted (black dotted lines) and detached flakes (grey line). a + b: difference in the lateral axis, c: difference in length, d: distance between points of percussion between predicted and detached flakes. (*b*) Mean differences between predicted and detached flakes for the three different expertise groups (error bars: +1s.d.). Black bars, expert; grey bars, intermediate; white bars, novice; asterisk, difference is significant at $p < 0.05$. Adapted from Nonaka *et al.* [56].

two percussive tasks, nut cracking and stone flaking, and we have also discovered behavioural correlates of experience and expertise in stone-knapping tasks. These may be informative of evolutionary divergences in the behavioural capacities of the chimpanzee and human lineages. Our behavioural results reviewed above can be seen as bridging the gap between work reviewed elsewhere in this issue on monkey tool use and social learning [85], and work on human tool use and on the evolution of increasing cognitive demands in hominin Palaeolithic stone tool traditions [86,87]. Such work does not directly address African ape–human cognitive and behavioural contrasts, nor do we yet have any brain imaging observations even for a human model of the circuits activated in a nut-cracking as compared with a simple stone-flaking task (see also [88,89]).

A number of potential anatomical contrasts have been hypothesized that could explain differences in human and chimpanzee tool-making skills, but some of these have not yet been experimentally validated. Stout & Chaminade [8] have identified activation of the posterior parietal area (PP, caudal intraparietal/

transverse occipital sulci) in experimental Oldowan stone flaking by modern humans. The PP is generally recognized as a site of major expansion and reorganization in humans [90,91] and may thus constitute a neural basis of what makes humans unique in tool-making skills [92]. However, because we do not have brain imaging data on areas that are activated in a nut-cracking task, we cannot be confident that the two tasks place qualitatively different demands on differing elements of cognitive systems (as opposed to quantitatively different demands on the same elements of the same cognitive systems), nor can we assume that PP contrasts found when comparing humans and monkeys will also be found when comparing humans with chimpanzees. At a more peripheral level, Walker [93] has recently proposed an explanation for the difference in strength in the human and chimpanzee limb systems (particularly the upper limb system) that may also have implications for skill in percussive task execution. He proposes that chimpanzees have fewer and larger motor units (systems of motoneurons signalling to the muscle fascicles to contract), which enables greater simultaneous force to be exerted.

He suggests that, in contrast, humans have a much greater range of motor units and more small units (fewer muscle fibres per nerve), enabling us to recruit muscles for more complex but less forceful tasks. If his conjecture is correct, then chimpanzees are optimized for strength in peak loading tasks (locomotion), while humans no longer need to exert the same degree of force in their upper limbs and have therefore evolved a capacity for finer and more rapidly varying forces, but with less maximal strength. Walker's hypothesis, however, has yet to be tested.

In contrast to such work, there is also substantial evidence of continuity between humans and other primates in many fundamental aspects of brain organization relevant to skilled tool use (see also [94]). Imaging studies show that primate tool use activates a distributed network of brain areas, localized in different but interconnected anatomical structures. Brain imaging studies in humans have shown that it is important to differentiate circuits involved in conceptual semantic knowledge about tool use (often based on tasks like pantomime, visual evaluation of a tool image, hearing tool manipulation, tool naming, etc.; e.g. [95–97]), and those involved in the selection and effective use of a functional tool to solve a mechanical problem [98,99]. Thus, human tool use activates a fronto-parietal praxis network involved in hand manipulation skills, as well as regions of the cerebellum and the basal ganglia [95], while macaque tool use also activates several cortical areas (intraparietal cortex, presupplementary motor area, premotor cortex) as well as the cerebellum and the basal ganglia [100]. Human tool use also activates a network more associated with conceptual aspects of tool use involving the left inferior frontal gyrus, left posterior middle temporal gyrus and bilateral fusiform cortex ([95], cf. [101]), for which comparative analyses of tool use in monkeys are presented elsewhere in this issue (cf. [85]).

Our own recent comparative anatomical research has focused on the coupled evolution of cortical and cerebellar circuits in primates, and their implications for the evolution of motor control systems. Motor control has been described as the 'Cinderella' of psychology [102], but research attention is increasing thanks to a convergence of interest between psychology and neuroscience, the rise of ecological psychology and of dynamical systems approaches and an enhanced awareness of the computational complexity of skilled movement among scientists programming action planning and execution routines in robotics [102]. In parallel with this, there has been an enormous increase in scientific research on the cerebellum, a structure that contains more than half of the neurons in the entire human brain and that has traditionally been seen as mainly concerned with motor learning and movement coordination, precision and timing, but which is now also thought to be involved in higher cognitive functions (e.g. [103]).

In an extension of recent evolutionary analyses of cortico-cerebellar systems in primate brains (e.g. [104–109]), our own recent work [110] has focused on the systemic relationships of the lateral cerebellum or 'cerebrocerebellum'. The cerebrocerebellum is

crucial for execution of prehensile upper limb movements, in both feed-forward and visual feedback-guided modes of regulation of ongoing movements in action execution. It works in conjunction with primary motor cortex and parietal association cortex (via the pontine nuclei) in the organization of skilled manual actions [103,109,111]. The cerebrocerebellum is also involved in a frontal cortical circuit (primarily prefrontal) via the basal ganglia, which is involved in novel motor sequence learning ('incremental acquisition of movements into a well-executed behaviour' [112, p. 252]). Imaging studies suggest that these circuits are conserved: for example, recent brain imaging work with macaques also shows activation of fronto-cerebellar and fronto-parietal circuits in tasks requiring extension of tool use to novel functional demands [113].

We have measured the volumes of the major cerebellar substructures in 19 living anthropoid primate species (including humans and chimpanzees), and have identified patterns of correlated and adaptive evolutionary size change in separate components of cortico-cerebellar systems within this diverse group (whose member species have not shared a common ancestor more recently than about 35 Myr BP) [110]. Our results indicate two main patterns of correlated evolution, indicative of selection acting repeatedly on integrated brain systems: one set of correlations involves the posterior neocortex, pons, cerebrocerebellum, dentate nucleus and thalamus, and the other set involves the frontal lobe, basal ganglia, cerebrocerebellum, dentate nucleus and thalamus (figure 10). We have suggested that these patterns of correlated evolution are specifically associated with selection acting to maintain the functional integrity of the two cortico-cerebellar circuits described above, and involved in the organization of skilled manual actions and in learning novel motor sequences. Our results therefore suggest that patterns of covarying size changes in neural systems involving profuse cortico-cerebellar connections are a major factor in explaining the evolution of anthropoid brain organization. We have also used phylogenetic comparative methods to reconstruct patterns of evolutionary divergence at successive phylogenetic branching events in the lineage leading to humans and chimpanzees. We infer that the *Homo-Pan* clade has come under strong positive selection for relative expansion of the frontal cortico-cerebellar system (with selection strongest in the human-specific branch). The marked expansion of the frontal cortico-cerebellar system in chimpanzees and humans is consistent with their increased social learning capacities, exemplified in their similar learning strategies of fine motor skills such as tool use.

Studies of human motor-skill learning demonstrate that the transitions from an initial effortful learning phase to more established performance levels involve increased activation of the cerebrocerebellum [114–117]. Motor adaptation (dynamic adjustment to environmental changes during execution of a learned motor sequence), in particular, is dependent on the intact cortico-cerebellar system that links cerebrocerebellar areas to parietal and motor cortex [118] and is involved in both kinematic and dynamic aspects of motor control [119]. Lesion studies have provided further evidence for the role of global cerebellar
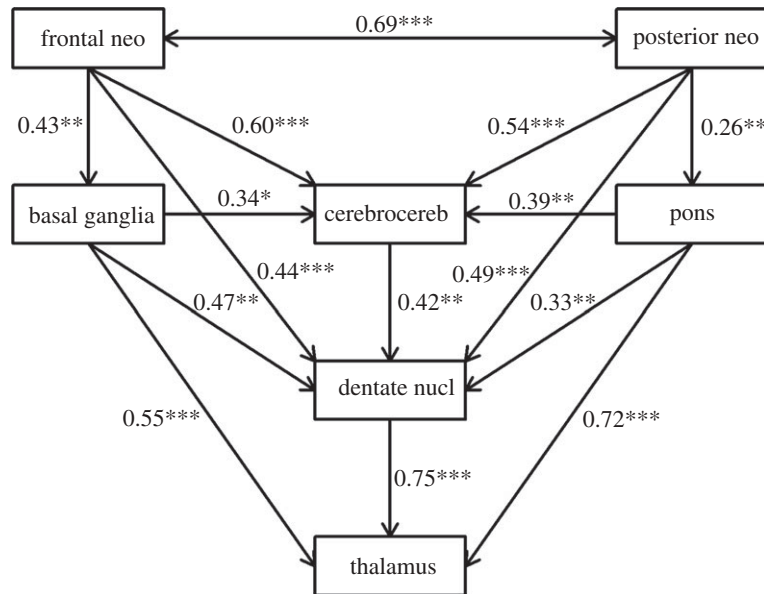
Figure 10. A summary of statistically significant bivariate correlations showing the concerted evolution of functionally inter-related structures across a sample of 19 anthropoid primate species' brains. Values indicate partial r^2 values in a phylogenetically controlled comparison. The correlations are significant at probability level: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. Each correlation represents the coevolutionary relation between two structures, partialling out the size of the rest of the brain (adapted from Smaers *et al.* [110]).

deficits in profound impairments of motor adaptation (e.g. [120]). Such studies, which focus on the neuropsychology of non-declarative, incremental motor-skill learning (where 'practice makes perfect') and on the acquisition of capacities for motor adaptation to varying environmental constraints (analogous to the varying hammer weights that we introduced in our behavioural studies of skilled percussive action), identify roles for brain systems and components of brain systems that complement the cortical mechanisms of action understanding and action planning discussed by other contributors to this volume. The main outcome of our comparative anatomical study, as summarized in this section, is its indirect confirmation of a fundamental evolutionary continuity in the organization of such cortico-cerebellar systems in anthropoid primates (including monkeys, apes and humans). We would expect such anatomical continuities to be reflected in underlying continuities of behavioural potential for motor-skill learning, although contrasts between species in absolute brain size and (if found) in finer grained aspects of brain architecture may affect individual species' aptitudes for learning and executing complex motor sequences in particular behavioural domains.

Our behavioural studies have also identified contrasts between the chimpanzee nut-cracking and the human stone-flaking tasks in functional parameters relating to understanding of the fracture mechanics of stone cores (which are much harder to learn to predict than are the fracture properties of hard casings of nuts). As discussed above, it is worth keeping in mind that chimpanzees (or bonobos such as Kanzi) somehow fall short in being able to visualize the properties of the core so as to exploit them to produce flakes [54,55]. A functional imaging study of brain regions activated in human subjects by physical reasoning in nut-cracking versus stone-flaking tasks would elucidate this problem, but to our knowledge, no such study has yet been conducted.

5. CONCLUSIONS

We have reviewed our experimental work on the nut-cracking and stone-flaking tasks, which has been designed within a dynamical systems framework and with reference to the ecological movement in psychology [12]. In this framework, the mastering of a technical skill depends on the capacity of an organism to set up the constraints of the system according to the task demand, and to mobilize adaptively the degrees of freedom of the system. At a behavioural level, the unfolding of the action may be viewed as an emergent process, at the interface of environmental opportunities available to the organism (affordances) and the set of constraints associated with the task. Nut cracking and stone knapping differ in task conditions because conchoidal fracture of a stone core requires more precise movement control, and an asymmetrical use of both hands, characterized by the simultaneous control of at least two variables (reciprocal orientation of the core and of the trajectory of the hammer, which keeps varying during the sequences of blows). We have also shown that one of the features characterizing expert stone knappers is the ability to predict accurately the effect of a percussive strike on fracture propagation in the core, which depends on properties of the core, such as external platform angle and core surface morphology, as well as properties of the executed strike, such as kinetic energy at impact and distance of the point of impact from the platform edge.

The neuroscience of such skilled movement, and of the capacity for adaptive response to varying functional opportunities (namely the affordances of the individual core), requires further investigation (cf. [87]). In particular, it would be helpful to compare patterns of brain activation in nut cracking and stone flaking in a cohort trained in the skilful execution of both tasks. An integrated approach to the concerted evolution of systems of functionally interrelated brain structures is

increasingly common in comparative neuroethological research. We expect that skilful task execution will depend on subcortical as well as cortical circuits, and we have discussed the evolutionary anatomy of the cortico-cerebellar systems as an appropriate focus for further comparative investigation.

REFERENCES

1 Roche, H. 2005 From simple flaking to shaping: stone-knapping evolution among early hominins. In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 35–48. Cambridge, UK: McDonald Institute for Archaeological Research.

2 Semaw, S. 2000 The world's oldest stone artefacts from Gona, Ethiopia: their implications for understanding stone technology and patterns of human evolution between 2.6–1.5 million years ago. *J. Archaeol. Sci.* **27**, 1197–1214. (doi:10.1006/jasc.1999.0592)

3 McPherron, S. P., Alemseged, Z., Marean, C. W., Wynn, J. G., Reed, D., Geraads, D., Bobe, R. & Béarat, H. A. 2010 Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* **466**, 857–860. (doi:10.1038/nature09248)

4 Panger, M. A., Brooks, A., Richmond, B. & Wood, B. 2002 Older than the Oldowan? Rethinking the emergence of hominin tool use. *Evol. Anthropol.* **11**, 235–245. (doi:10.1002/evan.10094)

5 Ambrose, S. H. 2001 Paleolithic technology and human evolution. *Science* **291**, 1748–1753. (doi:10.1126/science.1059487)

6 Roche, H., Delagnes, A., Grugal, J.-P., Feibel, C., Kibunjia, M., Mourre, V. & Texier, J. 1999 Early hominid stone tool production and technical skill 2.34 Myr ago in West Turkana, Kenya. *Nature* **399**, 57–60. (doi:10.1038/19959)

7 Steele, J. 1999 Stone legacy of skilled hands. *Nature* **399**, 24–25. (doi:10.1038/19869)

8 Stout, D. & Chaminade, T. 2007 The evolutionary neuroscience of tool making. *Neuropsychologica* **45**, 1091–1100. (doi:10.1016/j.neuropsychologia.2006.09.014)

9 Susman, R. L. 1998 Hand function and tool behavior in early hominids. *J. Hum. Evol.* **35**, 23–46. (doi:10.1006/jhev.1998.0220)

10 Wood, B. 1997 The oldest whodunnit in the world. *Nature* **385**, 292–293. (doi:10.1038/385292a0)

11 Harlacker, L. 2006 Knowledge and know-how in the Oldowan: an experimental approach. In *Skilled production and social reproduction: aspects of traditional stone-tool technologies* (eds J. Apel & K. Knutsson). Uppsala, Sweden: Societas Archaeologica Upsaliensis.

12 Roux, V. & Bril, B. 2005 General introduction: a dynamic system framework for studying a uniquely hominin innovation. In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 1–18. Cambridge, UK: McDonald Institute.

13 Stout, D., Toth, N., Schick, K. D. & Chaminade, T. 2008 Neural correlates of early Stone Age toolmaking: technology, language and cognition in human evolution. *Phil. Trans. R. Soc. B* **363**, 1939–1949. (doi:10.1098/rstb.2008.0001)

14 Toth, N. & Schick, K. D. 1993 Early stone industries and inferences regarding language and cognition. In *Tools, language and cognition in human evolution* (eds K. R. Gibson & T. Ingold), pp. 346–362. Cambridge, UK: Cambridge University Press.

15 Goodall, J. 1964 Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature* **201**, 1264–1266. (doi:10.1038/2011264a0)

16 Beatty, H. 1951 A note on the behavior of the chimpanzee. *J. Mammal.* **32**, 118.

17 Struhsaker, T. & Hunkeler, P. 1971 Evidence of tool-using by chimpanzees in the Ivory Coast. *Folia Primatol.* **15**, 212–219. (doi:10.1159/000155380)

18 Rahm, U. 1971 L'emploi d'outils par les chimpanzes de l'ouest de la Cote d'Ivoire. *Terre et Vie* **25**, 506–509.

19 Boesch, C. 1978 Nouvelles observations sur les chimpanzes de la foret de Tai (Cote d'Ivoire). *Terre et Vie* **32**, 195–201.

20 Sugiyama, Y. & Koman, J. 1979 Tool-using and -making behavior in wild chimpanzees at Bossou, Guinea. *Primates* **20**, 513–524. (doi:10.1007/BF02373433)

21 Sugiyama, Y. 1981 Observations on the population dynamics and behavior of wild chimpanzees at Bossou, Guinea, in 1979–1980. *Primates* **22**, 435–444. (doi:10.1007/BF02381236)

22 Toth, N. & Schick, K. 2009 The Oldowan: the tool making of early hominins and chimpanzees compared. *Annu. Rev. Anthropol.* **38**, 289–305. (doi:10.1146/annurev-anthro-091908-164521)

23 Plummer, T. 2004 Flaked stones and old bones: biological and cultural evolution at the dawn of technology. *Am. J. Phys. Anthropol.* **39**(Suppl.), 118–164. (doi:10.1002/ajpa.20157)

24 de Heinzelin, J., Clark, J. D., White, T., Hart, W., Renne, P., WoldeGabriel, G., Beyene, Y. & Vrba, E. 1999 Environment and behavior of 2.5-million-year-old Bouri hominids. *Science* **284**, 625–629. (doi:10.1126/science.284.5414.625)

25 Domínguez-Rodrigo, M., Pickering, T. R. & Bunn, H. T. 2010 Configurational approach to identifying the earliest hominin butchers. *Proc. Natl Acad. Sci. USA* **107**, 20 929–20 934. (doi:10.1073/pnas.1013711107)

26 Boesch, C. & Boesch, H. 1993 Diversity of tool use and tool making in wild chimpanzees. In *The use of tools by human and non-human primates* (eds A. Berthelet & J. Chavaillon), pp. 158–168. Oxford, UK: Clarendon Press.

27 Davidson, I. & McGrew, W. C. 2005 Stone tools and the uniqueness of human culture. *J. R. Anthropol. Inst.* **11**, 793–817. (doi:10.1111/j.1467-9655.2005.00262.x)

28 Marchant, L. F. & McGrew, W. C. 2005 Percussive technology: chimpanzee baobab smashing and the evolutionary modelling of hominin knapping. In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 341–350. Cambridge, UK: McDonald Institute for Archaeological Research.

29 Bril, B., Rein, R., Nonaka, T., Wenban-Smith, F. & Dietrich, G. 2010 The role of expertise in tool use: skill differences in functional action adaptation to task constraints. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 825–839. (doi:10.1037/a0018171)

30 Biryukova, E. V. & Bril, B. 2008 Organization of goal-directed action at a high level of motor skill: the case of stone knapping in India. *Motor Control* **12**, 181–209.

31 Jennings, J. S. & MacMillan, N. H. 1986 A tough nut to crack. *J. Mater. Sci.* **21**, 1517–1524. (doi:10.1007/BF01114704)

32 Bril, B., Dietrich, G., Foucart, J. & Hirata, S. 2009 Tool use as a way to assess cognition: how do captive chimpanzees handle the weight of the hammer when cracking a nut? *Anim. Cogn.* **12**, 217–235. (doi:10.1007/s10071-008-0184-x)

33 Foucart, J., Bril, B., Hirata, S., Morimura, N., Houki, C., Ueno, Y. & Matsuzawa, T. 2005 A preliminary analysis of nut-cracking movements in a captive chimpanzee: adaptation to the properties of tools and nuts. In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 147–158. Cambridge, UK: McDonald Institute for Archaeological Research.

34 Schrauf, C., Huber, L. & Visalberghi, E. 2008 Do capuchin monkeys use weight to select hammer tools? *Anim. Cogn.* **11**, 413–422. (doi:10.1007/s10071-007-0131-2)

35 Williams, S. H., Wright, B. W., Troung, V. D., Daubert, C. R. & Vinyard, C. Y. 2005 Mechanical properties of foods used in experimental studies of primate masticatory function. *Am. J. Primatol.* **67**, 329–346. (doi:10.1002/ajp.20189)

36 Peters, C. R. 1987 Nut-like oil seeds: food for monkeys, chimpanzees, humans, and probably ape-men. *Am. J. Phys. Anthropol.* **73**, 333–363. (doi:10.1002/ajpa.1330730306)

37 Visalberghi, E., Sabbatini, G., Spagnoletti, N., Andrade, F. R. D., Ottoni, E., Izar, P. & Fragaszy, D. 2008 Physical properties of palm fruits processed with tools by wild bearded capuchins (*Cebus libidinosus*). *Am. J. Primatol.* **70**, 884–891. (doi:10.1002/ajp.20578)

38 Koya, O. A. 2006 Palm nut cracking under repeated impact load. *J. Appl. Sci.* **11**, 2471–2475.

39 Pouydebat, E., Gorce, P., Bels, V. & Coppens, Y. 2006 Substrate optimization in nut cracking by capuchin monkeys (*Cebus apella*). *Am. J. Primatol.* **68**, 1017–1024. (doi:10.1002/ajp.20291)

40 Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P. & Fragaszy, D. 2009 Selection of effective stone tools by wild bearded capuchin monkeys. *Curr. Biol.* **19**, 213–217. (doi:10.1016/j.cub.2008.11.064)

41 Izawa, K. & Mizuno, A. 1977 Palm-fruit cracking behaviour of wild black-capped capuchin (*Cebus apella*). *Primates* **18**, 773–792. (doi:10.1007/BF02382930)

42 Sakura, O. & Matsuzawa, T. 1991 Flexibility of wild chimpanzee nut-cracking behavior using stone hammers and anvils: an experimental analysis. *Ethology* **87**, 237–248. (doi:10.1111/j.1439-0310.1991.tb00249.x)

43 Mora, R. & de la Torre, I. 2005 Percussion tools in Olduvai Beds I and II (Tanzania): implications for early human activities. *J. Anthropol. Archaeol.* **24**, 179–192. (doi:10.1016/j.jaa.2004.12.001)

44 Goren-Inbar, N., Sharon, G., Melamed, Y. & Kislev, M. 2002 Nuts, nut cracking, and pitted stones at Gesher Benot Ya'aqov, Israel. *Proc. Natl Acad. Sci. USA* **99**, 2455–2460. (doi:10.1073/pnas.032570499)

45 Lee, R. 1979 *The !Kung San: men, women, and work, in a foraging society.* Cambridge, MA: Harvard University Press.

46 Yellen, J. E. 1977 *Archaeological approaches to the present. Models for reconstructing the past.* New York, NY: Academic Press.

47 Bock, J. 2005 What makes a competent adult forager? In *Hunter–gatherer childhoods* (eds B. Hewlett & M. Lamb), pp. 109–128. New York, NY: Aldine de Gruyter.

48 Koya, O. A. & Faborode, M. O. 2005 Mathematical modeling of palm nut cracking based on Hertz's theory. *J. Biosyst. Eng.* **95**, 405–412.

49 Luedtke, B. E. 1992 *An archaeologist's guide to chert and flint.* Los Angeles, CA: Institute of Archaeology, University of California.

50 Cotterell, B. & Kamminga, J. 1987 The formation of flakes. *Am. Antiquity* **52**, 675–708. (doi:10.2307/281378)

51 Cotterell, B., Kamminga, J. & Dickson, F. D. 1985 The essential mechanics of conchoidal flaking. *Int. J. Fract.* **20**, 205–221. (doi:10.1007/BF00125471)

52 Cotterell, B. & Kamminga, J. 1990 *Mechanics of pre-industrial technology.* Cambridge, UK: Cambridge University Press.

53 Pelegrin, J. 2005 Remarks about archaeological techniques and methods of knapping: elements of a cognitive approach to stone knapping. In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 23–34. Cambridge, UK: McDonald Institute for Archaeological Research.

54 Schick, K. D., Toth, N., Garufi, G., Savage-Rumbaugh, E. S., Rumbaugh, D. & Sevcik, R. 1999 Continuing investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *J. Archaeol. Sci.* **26**, 821–832. (doi:10.1006/jasc.1998.0350)

55 Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., Sevcik, R. A. & Rumbaugh, D. M. 1993 Pan the tool-maker: investigations into the stone tool-making and tool-using capabilities of a bonobo. *J. Archaeol. Sci.* **20**, 81–91. (doi:10.1006/jasc.1993.1006)

56 Nonaka, T., Bril, B. & Rein, R. 2010 How do stone knappers predict and control the outcome of flaking? Implications for understanding early stone tool technology. *J. Hum. Evol.* **59**, 155–167. (doi:10.1016/j.jhevol.2010.04.006)

57 Dibble, H. L. & Rezek, Z. 2009 Introducing a new experimental design for controlled studies of flake formation: results for exterior platform angle, platform depth, angle of blow, velocity, and force. *J. Archaeol. Sci.* **36**, 1945–1954. (doi:10.1016/j.jas.2009.05.004)

58 Shelley, P. H. 1990 Variation in lithic assemblages: an experiment. *J. Field Archaeol.* **17**, 187–193.

59 Andrefsky Jr, W. 1998 *Lithics: macroscopic approaches to analysis.* Cambridge, UK: Cambridge University Press.

60 Inizan, M.-L., Reduron-Ballinger, M., Roche, H. & Tixier, J. 1999 *Technology and terminology of knapped stone.* Nanterre, Paris: Cercle de Recherches et d'Etudes Préhistoriques.

61 Bril, B. & Foucart, J. 2005 Enacting the perception of the affordances of potential tools. I: The case of children hammering. In *Studies in perception and action VIII* (eds H. Heft & K. L. Marsh), pp. 3–6. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

62 Bril, B., Roux, V. & Dietrich, G. 2005 Stone knapping: Khambhat (India), a unique opportunity? In *Stone knapping: the necessary conditions for a uniquely hominin behavior* (eds V. Roux & B. Bril), pp. 53–71. Cambridge, UK: McDonald Institute for Archaeological Research.

63 Bril, B., Roux, V. & Dietrich, G. 2000 Habiletés impliquées dans la taille des perles en calcédoine. Caractéristiques motrices et cognitives d'une action située complexe. In *Cornaline de l'inde: Des Pratiques Techniques de Cambay Aux Techno-systÀmes de l'Indus* (ed. V. Roux), pp. 207–332. Paris, France: Editions de la MSH.

64 Foucart, J. 2006 Etude comparee des habiletes operatoires et motrices de l'homme et du chimpanze pour une utilisation d'outils transprimatique: le cassage de noix. PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris, France.

65 Roux, V., Bril, B. & Dietrich, G. 1995 Skills and learning difficulties involved in stone knapping. *World Archaeol.* **27**, 63–87. (doi:10.1080/00438243.1995.9980293)

66 Nonaka, T. & Bril, B. 2011 Nesting of asymmetric functions in skilled bimanual action: dynamics of hammering behavior of bead craftsmen. *Hum. Mov. Sci.* **81**. (doi:10.1016/j.humov.2010.08.013)

67 Bernstein, N. A. 1996 On dexterity and its development. In *Dexterity and its development* (eds M. L. Latash & M. T. Turvey), pp. 1–235. Hillsdale, NJ: Lawrence Erlbaum Associates.

68 Ericsson, K. A. & Lehmann, A. C. 1996 Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu. Rev. Psychol.* **47**, 273–305. (doi:10.1146/annurev.psych.47.1.273)

69 Biryukova, E. V., Roby-Brami, A., Frolov, A. A. & Mokhtari, M. 2000 Kinematics of human arm reconstructed from spatial tracking system recording. *J. Biomech.* **33**, 985–995. (doi:10.1016/S0021-9290(00)00040-3)

70 Kortlandt, A. 1986 The use of stone tools by wild-living chimpanzees and earliest hominids. *J. Hum. Evol.* **15**, 77–132. (doi:10.1016/S0047-2484(86)80068-9)

71 Visalberghi, E., Fragaszy, D., Ottoni, E., Izar, P., de Oliveira, M. G. & Andrade, F. R. D. 2007 Characteristics of hammer stones and anvils used by wild bearded capuchin monkeys (*Cebus libidinosus*) to crack open palm nuts. *Am. J. Phys. Anthropol.* **132**, 426–444. (doi:10.1002/ajpa.20546)

72 Bril, B., Dietrich, G., Byriukova, L., Roby-Brami, A. & Roux, V. 2001 Hammering, adaptation to tool properties and raw material. *Poster presentation. International Workshop Knapping stone, a uniquely hominid behavior?, Abbaye des Premontrés, Pont-à-Mousson, 21–24 novembre.*

73 Kim, S. G., Ashe, J., Hendrich, K., Ellermann, J. M., Merkle, H., Ugurbil, K. & Georgopoulos, A. P. 1993 Functional magnetic resonance imaging of motor cortex: hemispheric asymmetry and handedness. *Science* **261**, 615–617. (doi:10.1126/science.8342027)

74 McManus, I. C. 1985 Handedness, language dominance and aphasia: a genetic model. *Psychol. Med.* **8**, 1–40.

75 Sainburg, R. L. 2000 Differences in control of limb dynamics during dominant and nondominant arm reaching. *J. Neurophysiol.* **83**, 2661–2675.

76 Guiard, Y. 1987 Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *J. Motor Behav.* **19**, 486–517.

77 Steele, J. & Uomini, N. 2009 Can the archaeology of manual specialization tell us anything about language evolution? A survey of the state of play. *Camb. Archaeol. J.* **19**, 97–110. (doi:10.1017/S0959774309000067)

78 Guiard, Y. 1987 Precursors to what? Theory is wanted for handedness in humans. Open peer commentary on 'Primate handedness reconsidered', by P. F. MacNeilage, M. G. Studdert-Kennedy & B. Lindblom. *Behav. Brain Sci.* **10**, 276–277. (doi:10.1017/S0140525X00047816)

79 Carson, R. G. 1993 Manual asymmetries: old problems and new directions. *Hum. Mov. Sci.* **12**, 479–506. (doi:10.1016/0167-9457(93)90001-6)

80 Sainburg, R. L. & Eckhardt, R. B. 2005 Optimization through lateralization: the evolution of handedness. *Behav. Brain Sci.* **28**, 611–612. (doi:10.1017/S0140525X05440108)

81 Wang, J. & Sainburg, R. L. 2007 The dominant and nondominant arms are specialized for stabilizing different features of task performance. *Exp. Brain Res.* **178**, 565–570. (doi:10.1007/s00221-007-0936-x)

82 Hopkins, W. D. & Rabinowitz, D. M. 1997 Manual specialisation and tool use in captive chimpanzees (*Pan troglodytes*): the effect of unimanual and bimanual strategies on hand preference. *Laterality* **2**, 267–277. (doi:10.1080/135765097397486)

83 Hopkins, W. D., Fernanadez-Carriba, S., Wesley, M. J., Hostetter, A., Pilcher, D. & Poss, S. 2001 The use of bouts and frequencies in the evaluation of hand preferences for a coordinated bimanual task in chimpanzees (*Pan troglodytes*): an empirical study comparing two different indices of laterality. *J. Comp. Psychol.* **115**, 294–299. (doi:10.1037/0735-7036.115.3.294)

84 Hopkins, W. D. 2006 Comparative and familial analysis of handedness in great apes. *Psychol. Bull.* **132**, 538–559. (doi:10.1037/0033-2909.132.4.538)

85 Macellini, S., Maranesi, M., Bonini, L., Simone, L., Rozzi, S., Ferrari, P. F. & Fogassi, L. 2012 Individual and social learning processes involved in the acquisition and generalization of tool use in macaques. *Phil. Trans. R. Soc. B* **367**, 24–36. (doi:10.1098/rstb.2011.0125)

86 Roby-Brami, A., Hermsdörfer, J., Roy, A. C. & Jacobs, S. 2012 A neuropsychological perspective on the link between language and praxis in modern humans. *Phil. Trans. R. Soc. B* **367**, 144–160. (doi:10.1098/rstb.2011.0122)

87 Stout, D. & Chaminade, T. 2012 Stone tools, language and the brain in human evolution. *Phil. Trans. R. Soc. B* **367**, 75–87. (doi:10.1098/rstb.2011.0099)

88 Frey, S. H. & Povinelli, D. J. 2012 Comparative investigations of manual action representations: evidence that chimpanzees represent the costs of potential future actions involving tools. *Phil. Trans. R. Soc. B* **367**, 48–58. (doi:10.1098/rstb.2011.0189)

89 Hopkins, W. D., Russell, J. L. & Schaeffer, J. A. 2012 The neural and cognitive correlates of aimed throwing in chimpanzees: a magnetic resonance image and behavioural study on a unique form of social tool use. *Phil. Trans. R. Soc. B.* **367**, 37–47. (doi:10.1098/rstb.2011.0195)

90 Orban, G. A., Van Essen, D. & Vanduffel, W. 2004 Comparative mapping of higher visual areas in monkeys and humans. *Trends Cogn. Sci.* **8**, 315–324. (doi:10.1016/j.tics.2004.05.009)

91 Zilles, K. 2005 Evolution of the human brain and comparative cyto- and receptor architecture. In *From monkey brain to human brain: a Fyssen Foundation symposium* (eds S. Dehaene, J. R. Duhamel, M. D. Hauser & G. Rizzolatti), pp. 41–56. Cambridge, MA: MIT Press.

92 Orban, G. A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J. T., Wardak, C., Durand, J. B. & Vanduffel, W. 2006 Mapping the parietal cortex of human and non-human primates. *Neuropsychologia* **44**, 2647–2667. (doi:10.1016/j.neuropsychologia.2005.11.001)

93 Walker, A. 2009 The strength of apes and the speed of humans. *Curr. Anthropol.* **50**, 229–234. (doi:10.1086/592023)

94 Hopkins, W. D. 2010 The comparative neuropsychology of tool use in primates with specific reference to chimpanzees and capuchin monkeys. In *Primate neuroethology* (eds M. Platt & A. Ghazanfar), pp. 587–614. Oxford, UK: Oxford University Press.

95 Lewis, J. W. 2006 Cortical networks related to human use of tools. *Neuroscience* **2**, 211–231.

96 Ramayya, A. G., Glasser, M. F. & Rilling, J. K. 2010 A DTI investigation of neural substrates supporting tool use. *Cereb. Cortex* **20**, 507–516. (doi:10.1093/cercor/bhp141)

97 Johnson-Frey, S. H., Newman-Norlund, R. & Crafton, S. T. 2005 A distributed left hemisphere network active during planning of everyday tool use skills. *Cereb. Cortex* **15**, 681–695. (doi:10.1093/cercor/bhh169)

98 Goldenberg, G. & Hagmann, S. 1998 Tool use and mechanical problem solving in apraxia. *Neuropsychologia* **36**, 581–589. (doi:10.1016/S0028-3932(97)00165-6)

99 Goldenberg, G. & Spatt, J. 2009 The neural basis of tool use. *Brain* **132**, 1645–1655. (doi:10.1093/brain/awp080)

100 Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., Onoe, H. & Iriki, A. 2001 Functional brain mapping of monkey tool use. *Neuroimage* **14**, 853–861. (doi:10.1006/nimg.2001.0878)

101 Johnson-Frey, S. H. 2004 The neural bases of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78. (doi:10.1016/j.tics.2003.12.002)

102 Rosenbaum, D. A. 2005 The Cinderella of psychology: the neglect of motor control in the science of mental life and behaviour. *Am. Psychol.* **60**, 308–317. (doi:10.1037/0003-066X.60.4.308)

103 Ramnani, N. 2006 The primate cortico-cerebellar system: anatomy and function. *Nat. Rev. Neurosci.* **7**, 511–522. (doi:10.1038/nrn1953)

104 Cantalupo, C. & Hopkins, W. D. 2010 The cerebellum and its contribution to complex tasks in higher primates: a comparative perspective. *Cortex* **46**, 821–830. (doi:10.1016/j.cortex.2009.10.004)

105 Herculano-Houzel, S. 2010 Coordinated scaling of cortical and cerebellar numbers of neurons. *Front. Neuroanat.* **4**, 1–8.

106 MacLeod, C. E., Zilles, K., Schleicher, A., Rilling, J. K. & Gibson, K. R. 2003 Expansion of the neocerebellum in Hominoidea. *J. Hum. Evol.* **44**, 401–429. (doi:10.1016/S0047-2484(03)00028-9)

107 Rilling, J. K. & Insel, T. R. 1998 Evolution of the primate cerebellum: differences in relative volume among monkeys, apes and humans. *Brain Behav. Evol.* **52**, 308–314. (doi:10.1159/000006575)

108 Weaver, A. H. 2005 Reciprocal evolution of the cerebellum and neocortex in fossil humans. *Proc. Natl Acad. Sci. USA* **102**, 3576–3580. (doi:10.1073/pnas.0500692102)

109 Whiting, B. A. & Barton, R. A. 2003 The evolution of the cortico-cerebellar complex in primates: anatomical connections predict patterns of correlated evolution. *J. Hum. Evol.* **44**, 3–10. (doi:10.1016/S0047-2484(02)00162-8)

110 Smaers, J. B., Steele, J. & Zilles, K. 2011 Modeling the evolution of cortico-cerebellar systems in primates. *Ann. NY Acad. Sci.* **1225**, 176–190. (doi:10.1111/j.1749-6632.2011.06003.x)

111 Middleton, F. A. & Strick, P. L. 2000 Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res. Rev.* **31**, 236–250. (doi:10.1016/S0165-0173(99)00040-5)

112 Doyon, J., Penhune, V. & Ungerleider, L. G. 2003 Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* **41**, 252–262. (doi:10.1016/S0028-3932(02)00158-6)

113 Obayashi, S., Suhara, T., Nagai, Y., Maeda, J., Hihara, S. & Iriki, A. 2002 Macaque prefrontal activity associated with extensive tool use. *Neuroreport* **13**, 2349–2354. (doi:10.1097/00001756-200212030-00036)

114 Albus, J. S. 1971 A theory of cerebellar function. *Math. Biosci.* **10**, 25–61. (doi:10.1016/0025-5564(71)90051-4)

115 Marr, D. 1969 A theory of cerebellar cortex. *J. Physiol. Lond.* **202**, 437–470.

116 Martin, T. A., Keating, J. G., Goodkin, H. P., Bastian, A. J. & Thach, W. T. 1996 Throwing while looking through prisms. 2. Specificity and storage of multiple gaze-throw calibrations. *Brain* **119**, 1199–1211. (doi:10.1093/brain/119.4.1199)

117 Puttemans, V., Wenderoth, N. & Swinnen, S. P. 2005 Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *J. Neurosci.* **25**, 4270–4278. (doi:10.1523/JNEUROSCI.3866-04.2005)

118 Doyon, J. & Benali, H. 2005 Reorganization and plasticity in the adult brain during learning of motor skills. *Curr. Opin. Neurobiol.* **15**, 161–167. (doi:10.1016/j.conb.2005.03.004)

119 Diedrichsen, J., Hashambhoy, Y., Rane, T. & Shadmehr, R. 2005 Neural correlates of reach errors. *J. Neurosci.* **25**, 9919–9931.

120 Maschke, M., Gomez, C. M., Ebner, T. J. & Konczak, J. 2004 Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *J. Neurophysiol.* **91**, 230–238. (doi:10.1152/jn.00557.2003)

Review

Stone tools, language and the brain in human evolution

Dietrich Stout[1],* and Thierry Chaminade[2]

[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 inferior 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

* Author for correspondence (dwstout@emory.edu).

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, somatosensory 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 redescribed 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
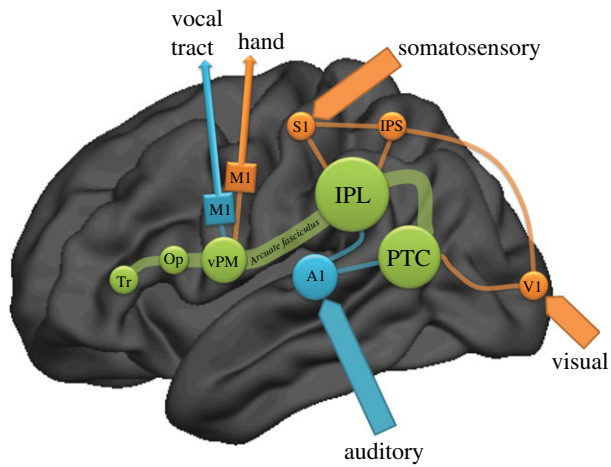
Figure 1. Cortical networks involved in speech and tool use. Green areas (posterior temporal cortex, PTC, inferior parietal lobule, IPL, and ventral premotor cortex, vPM, which, together with inferior frontal gyrus *pars opercularis*, Op, and *pars triangularis*, Tr, form Broca's area *sensu lato*) are proposed to participate similarly in speech and tool use. Blue areas are specific to speech perception (primary auditory cortices, A1) and execution (vocal tract primary motor cortex, M1). Orange areas are specific to tool use, including visual perception (primary visual area, V1), somatosensory perception (primary somatosensory cortex, S1), visuomotor integration in the intraparietal sulcus (IPS, part of the dorsal stream), as well as a primary motor cortex region involved in hand actions (M1).

[51] in suggesting that these complex structure–function relationships will be most profitably explored through an iterative research programme in which neuroscience data inform the fractionation of psychological processes (cf. [52]) and the fractionation of psychological processes motivates increasingly refined neuroscientific investigation.

These alternative interpretations of functional 'overlap' have important implications for our understanding of brain structure, function and evolution. However, all of them are at least theoretically consistent with some form of evolutionary interaction between the structures and functions in question. This includes the possibility that adjacent and functionally similar, but nevertheless distinct, adult structures could arise through evolutionary and ontogenetic differentiation from a common precursor (e.g. [6]) as well as the more obvious potential for behavioural co-optation of truly pluripotent (multifunctional) structures. Both possibilities are consistent with current theoretical views on the interaction of structural duplication, differentiation and plasticity [53,54] with functional degeneracy, redundancy and pluripotency [55] in cortical evolution. Better understanding of the relevant structure–function relationships in modern humans (and other primates, although this is not a focus of the current review) is a key step towards identifying the actual evolutionary relationships, if any, between specific behaviours like toolmaking and speech.

(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

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

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 nonlinguistic 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 organization 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 cobble '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
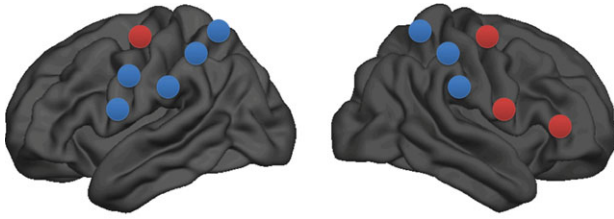
Figure 2. Parietal and prefrontal regions implicated in Oldowan and Acheulean toolmaking. Overlap between the two technologies is found in the inferior and superior parietal cortex as well as the intraparietal sulcus bilaterally, and in left ventral premotor cortex. Regions specifically involved in Acheulean toolmaking can be found in the frontal cortex, and in particular in the right hemisphere homologue of anterior Broca's area. Red circles, Acheulean only; blue circles, Oldowan and Acheulean.

record provides the first evidence of uniquely human capacities for manual praxis and these capacities can be specifically related to increased demands on an inferior parietal–ventral premotor circuit with important anatomical and computational similarities to that involved in phonological processing.

Such evidence cannot demonstrate an evolutionary connection but does corroborate and extend technological hypotheses of language origins by documenting a functional/anatomical link between a specific, archaeologically visible behaviour and a particular component of language competence. This leads to the suggestion [39] that selection acting on Oldowan toolmaking capacities could have favoured the elaboration of a praxic system that was subsequently co-opted to support the enhanced articulatory control required for speech. This proposal is broadly compatible with the evolutionary developmental scenario of Greenfield [6] and with Arbib's [29] MSH. It is distinguished from these hypotheses by its behavioural and chronological specificity and proposal that hominin adaptations for 'simple' individual praxis, not necessarily related to mirror system resonance, imitation or the complexity of abstract goal hierarchies, might also have contributed to producing a 'language-ready brain'.

(b) *Late Acheulean toolmaking*

Late Acheulean handaxe production activates the same dorsal stream structures implicated in Oldowan toolmaking, but with additional recruitment of right ventral premotor cortex and the dorsal portion of right IFG *pars triangularis* (figure 2). As described above (§2*d*), *pars triangularis* is associated with more abstract action representation and hierarchical organization, including semantic/syntactic integration. Recently, the dorsal portion of left *pars triangularis* has been specifically associated with working memory underpinning the ability to process sentences with long-distance structural separations between syntactically related elements [153]. This might be seen as analogous to the increased separation between functionally related technical actions seen in the relatively complex goal hierarchies of Late Acheulean toolmaking [47]. For example, the production of thin and symmetrical Late Acheulean handaxes requires highly

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 (§2*e*), 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
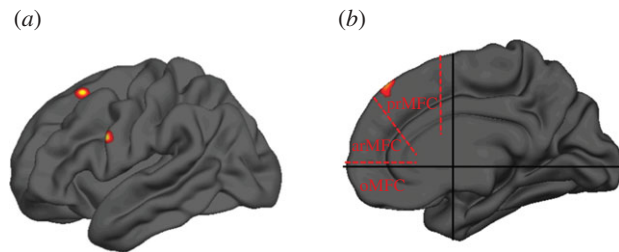
Figure 3. Brain activation during the observation of Lower Palaeolithic toolmaking. In technologically naive subjects (*a*), increased technological complexity (from Oldowan to Acheulean) is associated with increased motor resonance in left posterior inferior frontal gyrus. In expert subjects (*b*), increased technological complexity is instead associated with an increase in posterior rostral medial prefrontal activation reflecting the attribution of intention (medial frontal subdivisions in accordance with Vigneau *et al.* [133]: posterior rostral medial prefrontal cortex, prMFC; anterior rostral medial frontal cortex, arMFC, and orbital medial prefrontal cortex, oMFC).

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 'imitative' 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.

REFERENCES

1 Darwin, C. 2004 *The descent of man, and selection in relation to sex*. London, UK: Penguin Books.

2 Kimura, D. 1979 Neuromotor mechanisms in the evolution of human communication. In *Neurobiology of social communication in primates* (eds L. H. D. Steklis & M. J. Raleigh), pp. 179–219. New York, NY: Academic Press.

3 Lashley, K. 1951 The problem of serial order in behavior. In *Cerebral mechanisms in behavior* (ed. L. A. Jeffress), pp. 112–136. New York, NY: John Wiley.

4 Holloway, R. 1969 Culture: a human domain. *Curr. Anthropol.* **10**, 395–412. (doi:10.1086/201036)

5 Reynolds, P. C. 1976 Language and skilled activity. *Ann. NY Acad. Sci.* **280**, 150–166. (doi:10.1111/j.1749-6632.1976.tb25483.x)

6 Greenfield, P. M. 1991 Language, tools, and brain: the development and evolution of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**, 531–595. (doi:10.1017/S0140525X00071235)

7 Engels, F. 2003 The part played by labour in the transition from ape to man. In *Philosophy of technology* (eds R. C. Scharff & V. Dusek), pp. 71–77. London, UK: Blackwell.

8 Montagu, A. 1976 Toolmaking, hunting and the origin of language. *Ann. NY Acad. Sci.* **280**, 266–274. (doi:10.1111/j.1749-6632.1976.tb25493.x)

9 Isaac, G. L. 1976 Stages of cultural elaboration in the Pleistocene: possible archaeological indicators of the development of language capabilities. *Ann. NY Acad. Sci.* **280**, 275–288. (doi:10.1111/j.1749-6632.1976.tb25494.x)

10 Reynolds, P. C. 1993 The complementation theory of language and tool use. In *Tools, language and cognition in human evolution* (eds K. R. Gibson & T. Ingold), pp. 407–428. Cambridge, UK: Cambridge University Press.

11 Bookheimer, S. 2002 Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* **25**, 151–188. (doi:10.1146/annurev.neuro.25.112701.142946)

12 Higuchi, S., Chaminade, T., Imamizu, H. & Kawato, M. 2009 Shared neural correlates for language and tool use in Broca's area. *Neuroreport* **20**, 1376–1381. (doi:10.1097/WNR.0b013e3283315570)

13 Fink, G. R., Manjaly, Z. M., Stephan, K. E., Gurd, J. M., Zilles, K., Amunts, K. & Marshall, J. C. 2006 A role for Broca's area beyond language processing: evidence from neuropsychology and fMRI. In *Broca's region* (eds Y. Godzinsky & K. Amunts), pp. 254–268. Oxford, UK: Oxford University Press.

14 Passingham, R. E., Stephan, K. E. & Kotter, R. 2002 The antomical basis of functional localization in the cortex. *Nat. Rev. Neurosci.* **3**, 606–616.

15 Deacon, T. W. 1997 *The symbolic species: the co-evolution of language and the brain.* New York, NY: W.W. Norton.

16 Price, C. J. & Friston, K. J. 2002 Degeneracy and cognitive anatomy. *Trends Cogn. Sci.* **6**, 416–421. (doi:10.1016/S1364-6613(02)01976-9)

17 Gould, S. J. & Vrba, E. S. 1982 Exaptation: a missing term in the science of form. *Paleobiology* **8**, 4–15.

18 Broca, P. 2006 Comments regarding the seat of the faculty of spoken language, followed by an observation of aphemia (loss of speech). In *Broca's region* (eds Y. Grodzinsky & K. Amunts), pp. 291–304. Oxford, UK: Oxford University Press.

19 Hagoort, P. 2005 On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* **9**, 416–423. (doi:10.1016/j.tics.2005.07.004)

20 Pulvermüller, F. & Fadiga, L. 2010 Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* **11**, 351–360. (doi:10.1038/nrn2811)

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

22 Schubotz, R. I. & von Cramon, D. Y. 2003 Functional-anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. *NeuroImage* **20**, S120–S131. (doi:10.1016/j.neuroimage.2003.09.014)

23 Koechlin, E. & Jubault, T. 2006 Broca's area and the hierarchical organization of human behavior. *Neuron* **50**, 963–974. (doi:10.1016/j.neuron.2006.05.017)

24 Rizzolatti, G. & Craighero, L. 2004 The mirror-neuron system. *Ann. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)

25 Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)

26 Gallese, V. & Goldman, A. 1998 Mirror neurons and the simulation theory of mind-reading. *Trends Cogn. Sci.* **2**, 493–501. (doi:10.1016/S1364-6613(98)01262-5)

27 Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. 1967 Perception of the speech code. *Psychol. Rev.* **74**, 431–461. (doi:10.1037/h0020279)

28 Rizzolatti, G. & Arbib, M. A. 1998 Language within our grasp. *Trends Cogn. Sci.* **21**, 188–194.

29 Arbib, M. A. 2005 From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**, 105–124.

30 Hewes, G. W. 1973 Primate communication and the gestural origins of language. *Curr. Anthropol.* **14**, 5–24. (doi:10.1086/201401)

31 Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S. & Grafton, S. T. 2003 Actions or hand–object interactions? Human inferior frontal cortex and action observation. *Neuron* **39**, 1053–1058. (doi:10.1016/S0896-6273(03)00524-5)

32 Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W. & Boesch, C. 1999 Cultures in chimpanzees. *Nature* **399**, 682–685. (doi:10.1038/21415)

33 Zukow-Goldring, P. & Arbib, M. A. 2007 Affordances, effectivities, and assisted imitation: caregivers and the directing of attention. *Neurocomputing* **70**, 2181–2193. (doi:10.1016/j.neucom.2006.02.029)

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

World Archaeol. **26**, 158–171. (doi:10.1080/00438243.1994.9980270)

35 Wynn, T. 1995 Handaxe enigmas. *World Archaeol.* **27**, 10–24. (doi:10.1080/00438243.1995.9980290)

36 Chase, P. G. 1991 Symbols and Paleolithic artifacts: style, standardization, and the imposition of arbitrary form. *J. Anthropol. Archaeol.* **10**, 193–214. (doi:10.1016/0278-4165(91)90013-N)

37 Noble, W. & Davidson, I. 1996 *Human evolution, language and mind.* Cambridge, UK: Cambridge University Press.

38 Holloway, R. 1981 Culture, symbols and human brain evolution: a synthesis. *Dialect. Anthropol.* **5**, 287–303. (doi:10.1007/BF00246207)

39 Stout, D. & Chaminade, T. 2009 Making tools and making sense: complex, intentional behaviour in human evolution. *Camb. Archaeol. J.* **19**, 85–96. (doi:10.1017/S0959774309000055)

40 Gallese, V., Rochat, M., Cossu, G. & Sinigaglia, C. 2009 Motor cognition and its role in the phylogeny and ontogeny of action understanding. *Dev. Psychol.* **45**, 103–113. (doi:10.1037/a0014436)

41 Saxe, R. 2005 Against simulation: the argument from error. *Trends Cogn. Sci.* **9**, 174–179. (doi:10.1016/j.tics.2005.01.012)

42 Jacob, P. & Jeannerod, M. 2005 The motor theory of social cognition: a critique. *Trends Cogn. Sci.* **9**, 21–25. (doi:10.1016/j.tics.2004.11.003)

43 Arbib, M. A., Liebal, K. & Pika, S. 2008 Primate vocalization, gesture, and the evolution of human language. *Curr. Anthropol.* **49**, 1053–1076. (doi:10.1086/593015)

44 Stout, D. & Chaminade, T. 2007 The evolutionary neuroscience of tool making. *Neuropsychologia* **45**, 1091–1100. (doi:10.1016/j.neuropsychologia.2006.09.014)

45 Stout, D., Toth, N., Schick, K. D. & Chaminade, T. 2008 Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution. *Phil. Trans. R. Soc. B* **363**, 1939–1949. (doi:10.1098/rstb.2008.0001)

46 Faisal, A., Stout, D., Apel, J. & Bradley, B. 2010 The manipulative complexity of Lower Paleolithic stone toolmaking. *PLoS ONE* **5**, e13718. (doi:10.1371/journal.pone.0013718)

47 Stout, D. 2011 Stone toolmaking and the evolution of human culture and cognition. *Phil. Trans. R. Soc. B* **366**, 1050–1059. (doi:10.1098/rstb.2010.0369)

48 Stout, D., Passingham, R., Frith, C., Apel, J. & Chaminade, T. 2011 Technology, expertise and social cognition in human evolution. *Eur. J. Neurosci.* **33**, 1328–1338. (doi:10.1111/j.1460-9568.2011.07619.x)

49 Petrides, M., Cadoret, G. & Mackey, S. 2005 Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* **435**, 1235–1238. (doi:10.1038/nature03628)

50 Edelman, G. M. 1993 Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron* **10**, 115–125. (doi:10.1016/0896-6273(93)90304-A)

51 Adolphs, R. 2010 Conceptual challenges and directions for social neuroscience. *Neuron* **65**, 752–767. (doi:10.1016/j.neuron.2010.03.006)

52 Stout, D. 2010 The evolution of cognitive control. *Topics Cogn. Sci.* **2**, 614–630. (doi:10.1111/j.1756-8765.2009.01078.x)

53 Rakic, P. 2009 Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10**, 724–735. (doi:10.1038/nrn2719)

54 Krubitzer, L. 2009 In search of a unifying theory of complex brain evolution. *Ann. NY Acad. Sci.* **1156**, 44–67. (doi:10.1111/j.1749-6632.2009.04421.x)

55 Friston, K. J. & Price, C. J. 2003 Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* **7**, 151–152. (doi:10.1016/S1364-6613(03)00054-8)

56 Milner, A. D. & Goodale, M. A. 1995 *The visual brain in action.* Oxford, UK: Oxford University Press.

57 Frey, S. H. 2007 What puts the how in where? Tool use and the divided visual streams hypothesis. *Cortex* **43**, 368–375. (doi:10.1016/S0010-9452(08)70462-3)

58 Ramayya, A. G., Glasser, M. F. & Rilling, J. K. 2009 A DTI investigation of neural substrates supporting tool use. *Cereb. Cortex* **20**, 507–516. (doi:10.1093/cercor/bhp141)

59 Valyear, K. F., Cavina-Pratesi, C., Stiglick, A. J. & Culham, J. C. 2007 Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? *NeuroImage* **36**(Suppl. 2), T94–T108. (doi:10.1016/j.neuroimage.2007.03.031)

60 Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G. & Orban, G. A. 2009 The representation of tool use in humans and monkeys: common and uniquely human features. *J. Neurosci.* **29**, 11 523–11 539. (doi:10.1523/JNEUROSCI.2040-09.2009)

61 Jacobsen, C. F. 1934 Influence of motor and premotor lesions upon the retention of skilled movements in monkeys and chimpanzees. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* **13**, 225–247.

62 Schütz-Bosbach, S. & Prinz, W. 2007 Perceptual resonance: action-induced modulation of perception. *Trends Cogn. Sci.* **11**, 349–355. (doi:10.1016/j.tics.2007.06.005)

63 Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)

64 Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. & Matelli, M. 1988 Functional organization of inferior area 6 in the macaque monkey. *Exp. Brain Res.* **71**, 491–507. (doi:10.1007/BF00248742)

65 Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., Calzolari, F., Granieri, E. & Fadiga, L. 2009 Encoding of human action in Broca's area. *Brain* **132**, 1980–1988. (doi:10.1093/brain/awp118)

66 Binkofski, F., Buccino, G., Stephan, K., Rizzolatti, G., Seitz, R. J. & Freund, J. 1999 A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *Eur. J. Neurosci.* **11**, 3276–3286. (doi:10.1046/j.1460-9568.1999.00753.x)

67 Hickok, G. & Poeppel, D. 2004 Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Trends Cogn. Sci.* **4**, 131–138. (doi:10.1016/S1364-6613(00)01463-7)

68 Kelly, C., Uddin, L. Q., Shehzad, Z., Margulies, D. S., Castellanos, F. X., Milham, M. P. & Petrides, M. 2010 Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. *Eur. J. Neurosci.* **32**, 383–398. (doi:10.1111/j.1460-9568.2010.07279.x)

69 Xiang, H.-D., Fonteijn, H. M., Norris, D. G. & Hagoort, P. 2010 Topographical functional connectivity pattern in the perisylvian language networks. *Cereb. Cortex* **20**, 549–560. (doi:10.1093/cercor/bhp119)

70 Buccino, G. *et al.* 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404.

71 Simon, O., Mangin, J. F., Cohen, L., Le Bihan, D. & Dehaene, S. 2002 Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron* **33**, 475–487. (doi:10.1016/S0896-6273(02)00575-5)

72 Pa, J. & Hickok, G. 2008 A parietal-temporal sensory-motor integration area for the human vocal tract:

evidence from an fMRI study of skilled musicians. *Neuropsychologia* **46**, 362–368. (doi:10.1016/j.neuropsychologia.2007.06.024)

73 Price, C. J. 2010 The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann. NY Acad. Sci.* **1191**, 62–88. (doi:10.1111/j.1749-6632.2010.05444.x)

74 Rizzolatti, G., Luppino, G. & Matelli, M. 1998 The organization of the cortical motor system: new concepts. *Electroencephalogr. Clin. Neurophysiol.* **106**, 283–296. (doi:10.1016/S0013-4694(98)00022-4)

75 Emmorey, K., Mehta, S. & Grabowski, T. J. 2007 The neural correlates of sign versus word production. *NeuroImage* **36**, 202–208. (doi:10.1016/j.neuroimage.2007.02.040)

76 Buxbaum, L. J. 2001 Ideomotor apraxia: a call to action. *Neurocase* **7**, 445–458. (doi:10.1093/neucas/7.6.445)

77 Johnson-Frey, S. H. 2004 The neural bases of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78. (doi:10.1016/j.tics.2003.12.002)

78 Chao, L. L. & Martin, A. 2000 Representation of manipulable man-made objects in the dorsal stream. *NeuroImage* **12**, 478–484. (doi:10.1006/nimg.2000.0635)

79 Creem-Regehr, S. H. & Lee, J. N. 2005 Neural representations of graspable objects: are tools special? *Cogn. Brain Res.* **22**, 457–469. (doi:10.1016/j.cogbrainres.2004.10.006)

80 Chaminade, T., Meltzoff, A. N. & Decety, J. 2002 Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *NeuroImage* **15**, 318–328. (doi:10.1006/nimg.2001.0981)

81 Johnson-Frey, S. H., Newman-Norlund, R. & Grafton, S. T. 2005 A distributed left hemisphere network active during planning of everyday tool use skills. *Cereb. Cortex* **15**, 681–695. (doi:10.1093/cercor/bhh169)

82 Paulesu, E., Frith, C. & Frackowiak, R. 1993 The neural correlates of the verbal component of working memory. *Nature* **362**, 342–345. (doi:10.1038/362342a0)

83 Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S. & Dupoux, E. 2003 Phonological grammar shapes the auditory cortex: a functional magnetic resonance imaging study. *J. Neurosci.* **23**, 9541–9546.

84 Blakemore, S. J., Wolpert, D. M. & Frith, C. D. 1998 Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* **1**, 635–640. (doi:10.1038/2870)

85 Eickhoff, S. B., Amunts, K., Mohlberg, H. & Zilles, K. 2006 The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. *Cereb. Cortex* **16**, 268–279. (doi:10.1093/cercor/bhi106)

86 Kilner, J., Friston, K. & Frith, C. 2007 Predictive coding: an account of the mirror neuron system. *Cogn. Process.* **8**, 159–166. (doi:10.1007/s10339-007-0170-2)

87 Gazzola, V. & Keysers, C. 2009 The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex* **19**, 1239–1255. (doi:10.1093/cercor/bhn181)

88 Wolpert, D., Doya, K. & Kawato, M. 2003 A unifying computational framework for motor control and social interaction. *Phil. Trans. R. Soc. Lond. B* **358**, 593–602. (doi:10.1098/rstb.2002.1238)

89 Chaminade, T., Meltzoff, A. & Decety, J. 2005 An fMRI study of imitation: action representation and body schema. *Neuropsychologia* **43**, 115–127. (doi:10.1016/j.neuropsychologia.2004.04.026)

90 Buxbaum, L. J., Kyle, K. M. & Menon, R. 2005 On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cogn. Brain Res.* **25**, 226–239. (doi:10.1016/j.cogbrainres.2005.05.014)

91 Peschke, C., Ziegler, W., Kappes, J. & Baumgaertner, A. 2009 Auditory-motor integration during fast repetition: the neuronal correlates of shadowing. *NeuroImage* **47**, 392–402. (doi:10.1016/j.neuroimage.2009.03.061)

92 Iacoboni, M. & Wilson, S. M. 2006 Beyond a single area: motor control and language within a neural architecture encompassing Broca's area. *Cortex* **42**, 503. (doi:10.1016/S0010-9452(08)70387-3)

93 Shaywitz, B. A. et al. 1994 Localization of semantic processing using functional magnetic resonance imaging. *Human Brain Mapping* **2**, 149–158. (doi:10.1002/hbm.460020304)

94 Friederici, A. D. 2002 Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* **6**, 78–84. (doi:10.1016/S1364-6613(00)01839-8)

95 Beauchamp, M. S., Lee, K. E., Haxby, J. V. & Martin, A. 2002 Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* **34**, 149–159. (doi:10.1016/S0896-6273(02)00642-6)

96 Kellenbach, M. L., Brett, M. & Patterson, K. 2003 Actions speak louder than functions: the importance of manipulability and action in tool representation. *J. Cogn. Neurosci.* **15**, 30–45. (doi:10.1162/089892903321107800)

97 Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H. & Damasio, A. R. 2003 Neural correlates of conceptual knowledge for actions. *Cogn. Neuropsychol.* **20**, 409–432. (doi:10.1080/02643290244000248)

98 Chang, E. F., Rieger, J. W., Johnson, K., Berger, M. S., Barbaro, N. M. & Knight, R. T. 2010 Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.* **13**, 1428–1432.

99 Schultz, J., Friston, K. J., O'Doherty, J., Wolpert, D. M. & Frith, C. D. 2005 Activation in posterior superior temporal sulcus parallels parameter inducing the percept of animacy. *Neuron* **45**, 625–635. (doi:10.1016/j.neuron.2004.12.052)

100 Puce, A. & Perrett, D. 2003 Electrophysiology and brain imaging of biological motion. *Phil. Trans. R. Soc. Lond. B* **358**, 435. (doi:10.1098/rstb.2002.1221)

101 Beauchamp, M. S., Lee, K. E., Argall, B. D. & Martin, A. 2004 Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* **41**, 809–823. (doi:10.1016/S0896-6273(04)00070-4)

102 Pietrini, P., Furey, M. L., Ricciardi, E., Gobbini, M. I., Wu, W. H. C., Cohen, L., Guazzelli, M. & Haxby, J. V. 2004 Beyond sensory images: object-based representation in the human ventral pathway. *Proc. Natl Acad. Sci. USA* **101**, 5658. (doi:10.1073/pnas.0400707101)

103 Peuskens, H., Vanrie, J., Verfaillie, K. & Orban, G. 2005 Specificity of regions processing biological motion. *Eur. J. Neurosci.* **21**, 2864–2875. (doi:10.1111/j.1460-9568.2005.04106.x)

104 Tulving, E. 1972 Episodic memory and semantic memory. In *Organization of memory* (eds E. Tulving & W. Donaldson), pp. 381–402. New York, NY: Academic Press.

105 Chao, L. L., Haxby, J. V. & Martin, A. 1999 Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat. Neurosci.* **2**, 913–919. (doi:10.1038/13217)

106 Lewis, J. W. 2006 Cortical networks related to human use of tools. *Neuroscientist* **12**, 211–231. (doi:10.1177/1073858406288327)

107 Marshal, N., Faust, M., Hendler, T. & Jung-Beeman, M. 2009 An fMRI study of processing novel metaphorical sentences. *Laterality* **14**, 30–54.

108 Hurford, J. R. 2003 The neural basis of predicate-argument structure. *Behav. Brain Sci.* **26**, 261–283.

109 Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F. & Braun, A. R. 2009 Symbolic gestures and spoken language are processed by a common neural system. *Proc. Natl Acad. Sci. USA* **106**, 20 664–20 669. (doi:10.1073/pnas.0909197106)

110 Caramazza, A. & Mahon, B. Z. 2003 The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* **7**, 354. (doi:10.1016/S1364-6613(03)00159-1)

111 Patterson, K., Nestor, P. J. & Rogers, T. T. 2007 Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* **8**, 976–987. (doi:10.1038/nrn2277)

112 Hamzei, F., Rijntjes, M., Dettmers, C., Glauche, V., Weiller, C. & Buchel, C. 2003 The human action recognition system and its relationship to Broca's area: an fMRI study. *NeuroImage* **19**, 637–644. (doi:10.1016/S1053-8119(03)00087-9)

113 Badre, D. & Wagner, A. D. 2007 Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia* **45**, 2883–2901. (doi:10.1016/j.neuropsychologia.2007.06.015)

114 Race, E. A., Shanker, S. & Wagner, A. D. 2009 Neural priming in human frontal cortex: multiple forms of learning reduce demands on the prefrontal executive system. *J. Cogn. Neurosci.* **21**, 1766–1781. (doi:10.1162/jocn.2009.21132)

115 Shipp, S. 2005 The importance of being agranular: a comparative account of visual and motor cortex. *Phil. Trans. R. Soc. B* **360**, 797–814. (doi:10.1098/rstb.2005.1630)

116 Glasser, M. F. & Rilling, J. K. 2008 DTI tractography of the human brain's language pathways. *Cereb. Cortex* **18**, 2471–2482. (doi:10.1093/cercor/bhn011)

117 Anwander, A., Tittgemeyer, M., von Cramon, D. Y., Friederici, A. D. & Knösche, T. R. 2007 Connectivity-based parcellation of Broca's area. *Cereb. Cortex* **17**, 816–825. (doi:10.1093/cercor/bhk034)

118 Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G. & Orban, G. A. 2005 Observing others: multiple action representation in the frontal lobe. *Science* **310**, 332–336. (doi:10.1126/science.1115593)

119 Majdandzic, J., Bekkering, H., Van Schie, H. T. & Toni, I. 2009 Movement-specific repetition suppression in ventral and dorsal premotor cortex during action observation. *Cereb. Cortex* **19**, 2736. (doi:10.1093/cercor/bhp049)

120 Snyder, H. R., Feigenson, K. & Thompson-Schill, S. L. 2007 Prefrontal cortical response to conflict during semantic and phonological tasks. *J. Cogn. Neurosci.* **19**, 761–775. (doi:10.1162/jocn.2007.19.5.761)

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

122 Caspers, S., Zilles, K., Laird, A. R. & Eickhoff, S. B. 2010 ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage* **50**, 1148–1167. (doi:10.1016/j.neuroimage.2009.12.112)

123 Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B. & Tzourio-Mazoyer, N. 2006 Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage* **30**, 1414–1432. (doi:10.1016/j.neuroimage.2005.11.002)

124 Molnar-Szakacs, I., Kaplan, J., Greenfield, P. M. & Iacoboni, M. 2006 Observing complex action sequences: the role of the fronto-parietal mirror neuron system. *NeuroImage* **33**, 923–935. (doi:10.1016/j.neuroimage.2006.07.035)

125 Menenti, L., Petersson, K. M., Scheeringa, R. & Hagoort, P. 2009 When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. *J. Cogn. Neurosci.* **21**, 2358–2368. (doi:10.1162/jocn.2008.21163)

126 Ross, E. D. & Monnot, M. 2008 Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain Lang.* **104**, 51–74. (doi:10.1016/j.bandl.2007.04.007)

127 Hartmann, K., Goldenberg, G., Daumuller, M. & Hermsdorfer, J. 2005 It takes the whole brain to make a cup of coffee: the neuropsychology of naturalistic actions involving technical devices. *Neuropsychologia* **43**, 625–637. (doi:10.1016/j.neuropsychologia.2004.07.015)

128 Frey, S. H. & Gerry, V. 2006 Modulation of neural activity during observational learning of action and their sequential orders. *J. Neurosci.* **26**, 13 194–13 201. (doi:10.1523/JNEUROSCI.3914-06.2006)

129 Husain, M. & Nachev, P. 2007 Space and the parietal cortex. *Trends Cogn. Sci.* **11**, 30–36. (doi:10.1016/j.tics.2006.10.011)

130 Wildgruber, D., Pihan, H., Ackermann, H., Erb, M. & Grodd, W. 2002 Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence, and sex. *NeuroImage* **15**, 856–869. (doi:10.1006/nimg.2001.0998)

131 Jubault, T., Ody, C. & Koechlin, E. 2007 Serial organization of human behavior in the inferior parietal cortex. *J. Neurosci.* **27**, 11028–11036. (doi:10.1523/JNEUROSCI.1986-07.2007)

132 Hamilton, A. F. d. C. & Grafton, S. T. 2008 Action outcomes are represented in human inferior frontoparietal cortex. *Cereb. Cortex* **18**, 1160–1168. (doi:10.1093/cercor/bhm150)

133 Vigneau, M. *et al.* 2011 What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *NeuroImage* **54**, 577–593. (doi:10.1016/j.neuroimage.2010.07.036)

134 Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T. & McCarthy, G. 2005 Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cereb. Cortex* **15**, 1866. (doi:10.1093/cercor/bhi064)

135 Blakemore, S.-J. & Decety, J. 2001 From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* **2**, 561–567.

136 van der Linden, M., van Turennout, M. & Fernández, G. 2010 Category training induces cross-modal object representations in the adult human brain. *J. Cogn. Neurosci.* **23**, 1315–1331. (doi:10.1162/jocn.2010.21522)

137 Rota, G., Sitaram, R., Veit, R., Erb, M., Weiskopf, N., Dogil, G. & Birbaumer, N. 2009 Self regulation of regional cortical activity using real time fMRI: the right inferior frontal gyrus and linguistic processing. *Hum. Brain Mapp.* **30**, 1605–1614. (doi:10.1002/hbm.20621)

138 Aron, A. A., Robbins, T. W. & Poldrack, R. A. 2004 Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* **8**, 170–177. (doi:10.1016/j.tics.2004.02.010)

139 Verbruggen, F., Aron, A. R., Stevens, M. A. & Chambers, C. D. 2010 Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 13 966–13 971. (doi:10.1073/pnas.1001957107)

140 MacNeilage, P. F. 2008 *The origin of speech.* Oxford, UK: Oxford University Press.

141 Sergent, J. 1982 The cerebral balance of power: confrontation or cooperation? *J. Exp. Psychol. Hum. Percept. Perform.* **8**, 253. (doi:10.1037/0096-1523.8.2.253)

142 Iturria-Medina, Y. *et al.* 2011 Brain hemispheric structural efficiency and interconnectivity: rightward asymmetry in human and nonhuman primates. *Cereb. Cortex* **21**, 56–67. (doi:10.1093/cercor/bhq058)

143 MacNeilage, P. F., Studdert-Kennedy, M. G. & Lindblom, B. 1984 Functional precursors to language and its lateralization. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **246**, R912–R914.

144 Wynn, T. & Coolidge, F. L. 2004 The expert Neandertal mind. *J. Hum. Evol.* **46**, 467–487. (doi:10.1016/j.jhevol.2004.01.005)

145 Wadley, L., Hodgskiss, T. & Grant, M. 2009 Implications for complex cognition from the hafting of tools with compound adhesives in the Middle Stone Age, South Africa. *Proc. Natl Acad. Sci. USA* **106**, 9590–9594. (doi:10.1073/pnas.0900957106)

146 Semaw, S. *et al.* 2003 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *J. Hum. Evol.* **45**, 169–177. (doi:10.1016/S0047-2484(03)00093-9)

147 Nonaka, T., Bril, B. & Rein, R. 2010 How do stone knappers predict and control the outcome of flaking? Implications for understanding early stone tool technology. *J. Hum. Evol.* **59**, 155–167. (doi:10.1016/j.jhevol.2010.04.006)

148 Bril, B., Rein, R., Nonaka, T., Wenban-Smith, F. & Dietrich, G. 2010 The role of expertise in tool use: skill differences in functional action adaptations to task constraints. *J. Exp. Psychol. Hum. Percept. Perform.* **36**, 825–839. (doi:10.1037/a0018171)

149 Toth, N. 1985 The Oldowan reassessed: a close look at early stone artifacts. *J. Archaeol. Sci.* **12**, 101–120. (doi:10.1016/0305-4403(85)90056-1)

150 Bril, B. & Roux, V. 2005 Synthesis and speculations. In *Stone knapping: the necessary conditions for a uniquely hominin behaviour* (eds V. Roux & B. Bril), pp. 353–355. Cambridge, UK: McDonald Institute for Archaeological Research.

151 Wynn, T. & McGrew, W. 1989 An ape's view of the Oldowan. *Man* **24**, 383–398. (doi:10.2307/2802697)

152 Toth, N., Schick, K. D. & Semaw, S. 2006 A comparative study of the stone tool-making skills of *Pan*, *Australopithecus*, and *Homo sapiens*. In *The Oldowan: case studies into the earliest stone age* (eds N. Toth & K. D. Schick), pp. 155–222. Gosport, IN: Stone Age Institute Press.

153 Makuuchi, M., Bahlmann, J., Anwander, A. & Friederici, A. D. 2009 Segregating the core computational faculty of human language from working memory. *Proc. Natl Acad. Sci. USA* **106**, 8362–8367. (doi:10.1073/pnas.0810928106)

154 Bor, D., Duncan, J., Lee, A. C. H., Parr, A. & Owen, A. M. 2006 Frontal lobe involvement in spatial span: converging studies of normal and impaired function. *Neuropsychologia* **44**, 229–237. (doi:10.1016/j.neuropsychologia.2005.05.010)

155 Smith, E. E., Jonides, J. & Koeppe, R. A. 1996 Dissociating verbal and spatial working memory using PET. *Cereb. Cortex* **6**, 11. (doi:10.1093/cercor/6.1.11)

156 Csibra, G. & Gergely, G. 2011 Natural pedagogy as evolutionary adaptation. *Phil. Trans. R. Soc. B* **366**, 1149–1157. (doi:10.1098/rstb.2010.0319)

157 Amodio, D. M. & Frith, C. D. 2006 Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277. (doi:10.1038/nrn1884)

158 Holloway, R. L., Broadfield, D. & Yuan, M. 2004 *The human fossil record, volume 3, brain endocasts–the paleoneurological evidence*. Hoboken, NJ: Wiley-Liss.

159 Kaplan, H., Hill, K., Lancaster, J. & Hurtado, A. M. 2000 A theory of human life history evolution: diet, intelligence, and longevity. *Evol. Anthropol.* **9**, 156–185. (doi:10.1002/1520-6505(2000)9:4<156::AID-EVAN5>3.0.CO;2-7)

160 Tomasello, M. 1999 *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Research

Articulatory capacity of Neanderthals, a very recent and human-like fossil hominin

Anna Barney[1], Sandra Martelli[2,3], Antoine Serrurier[1] and James Steele[2,*]

[1]*Institute of Sound and Vibration Research, University of Southampton, University Road Southampton SO17 1BJ, UK*
[2]*AHRC Centre for the Evolution of Cultural Diversity, Institute of Archaeology, University College London, 31-34 Gordon Square, London WC1H 0PY, UK*
[3]*UCL Research Department of Cell and Developmental Biology, Gower Street, London WC1E 6BT, UK*

Scientists seek to use fossil and archaeological evidence to constrain models of the coevolution of human language and tool use. We focus on Neanderthals, for whom indirect evidence from tool use and ancient DNA appears consistent with an adaptation to complex vocal-auditory communication. We summarize existing arguments that the articulatory apparatus for speech had not yet come under intense positive selection pressure in Neanderthals, and we outline some recent evidence and analyses that challenge such arguments. We then provide new anatomical results from our own attempt to reconstruct vocal tract (VT) morphology in Neanderthals, and document our simulations of the acoustic and articulatory potential of this reconstructed Neanderthal VT. Our purpose in this paper is not to polarize debate about whether or not Neanderthals were human-like in all relevant respects, but to contribute to the development of methods that can be used to make further incremental advances in our understanding of the evolution of speech based on fossil and archaeological evidence.

Keywords: Neanderthal; vocal tract; hyoid; quantal vowels; evolution of speech

1. THE SPEECH, HANDEDNESS AND TOOL-USE NEXUS IN OUR CLOSEST EXTINCT RELATIVES: THE NEANDERTHALS

The evolution of the human cognitive systems that underlie praxis, tool use, language and speech, is the thematic focus of this Special Issue. Numerous attempts have been made to reconstruct the evolutionary trajectory leading to human language. In one recent example, Arbib [1,2] has proposed a 'mirror system' model in which the language system evolves from complex imitation of manual praxis (involving the capacity for social learning of longer sequences of novel and hierarchically organized actions), via a manual protosign stage (involving pantomime gestures by the signaller, with conventionalized gestures to disambiguate the meaning of these pantomimes), to a protolanguage stage in which vocal gestures accompany and 'invade' the communicative domain of these manual gestures. The last stage (fully grammatical linguistic structure) may then have been reached by cumulative cultural evolution (and not by genetic adaptation: [3,4]). Another

version of this model by Corballis [5,6] proposes that left-hemisphere lateralization of language processing (and human-handedness) may have been associated with the shift to the vocal channel in the sixth 'protolanguage' phase. Left-hemisphere lateralization of cortical vocal perception circuitry is held to be phylogenetically old and widespread in extant primates, so that as Corballis ([5], p. 197) puts it, 'language may have gone from hand to mouth, while lateralization went from mouth to hand'. In Corballis' [5,6] version of the model, the emergence of a bias towards right-handedness in the archaeological record is, therefore, a diagnostic marker of a vocal protolanguage.

At the time of writing we have archaeological and fossil anatomical evidence of population-level right-handedness in tool use in *Homo heidelbergensis*, Neanderthals and anatomically modern *Homo sapiens* [7–9]; these hominins are all relatively large-brained. We also have suggestive evidence of speech-relevant adaptations in the same three species from hyoid bone morphology [10,11], from analysis of the thoracic spinal canal [12], and from ancient DNA (the presence of the human form of *FOXP2* in Neanderthals: [13]). Thus, we have no *a priori* reason to doubt that Neanderthals had at least reached the vocal protolanguage stage.

However, the level of grammatical structure of Neanderthal vocal utterances remains speculative: we do not know whether or not the required biological and/or cultural preconditions were in place for the stable

cultural evolution of linguistic structure and usage as seen in human societies today. Some scientists have suggested that clues can be gleaned from the hierarchical and compositional organization of Neanderthal tool-using behaviour (cf. [14,15]); for instance, Ambrose [16] notes that by about 300 kyr BP, there was an acceleration of cultural evolution both in the populations ancestral to Neanderthals and in the African populations ancestral to early anatomically modern humans (MHs), including the appearance of composite tools (evidence of hafting of stone projectile points and other tools) requiring the assembly of at least three elements (the stone tool, the haft and the binding material). He suggests that such compositional assembly rules may be analogous to those of linguistic grammars. Others [17] have suggested that evidence of knot tying would provide a more exact analogy with (and marker for) the cognitive operations required for grammar. The evidence is, however, as yet too sparse to characterize the binding methods used by Neanderthals in tool hafting, and moreover, if fully grammatical languages emerged through cumulative cultural evolution, then such evidence could only indicate the presence of cognitive pre-adaptations. Further experimental analysis is required of the cognitive and behavioural organization of Neanderthal stone technologies (cf. the methodologies of Pastra & Aloimonos [18] and Stout & Chaminade [19]). In this paper, we shall focus instead on methods for assessing fossil evidence for the evolution of the Neanderthal vocal tract (VT), and thus of one aspect of Neanderthals' potential capacity for articulate speech.

(a) Hominin vocal tract morphology as evidence of selection for articulatory potential

Current work on fossil evidence for the evolution of fully grammatical speech must inevitably recognize the towering influence of the work of Philip Lieberman. Writing at a time when Chomsky's ideas were in the ascendant, and when cognitive science was flowering, Lieberman demanded that attention be given not just to syntax and semantics, but also to acoustic and articulatory phonology. In a *tour de force* of synthesis, he set out a new hypothesis about the Darwinian evolution of language under natural selection, proposing that he could identify adaptive and unique features of human VT anatomy, which were necessary for and diagnostic of spoken language. The evolution of these features could also be traced, Lieberman argued, in the hominin fossil record. The basic Lieberman hypothesis (e.g. papers reprinted in Lieberman [20]) states that speech perception characteristically requires listeners to decode a stream of varying, serially ordered acoustic output issuing at an exceptionally fast rate from the speaker's VT. In terms of syntactic and semantic content, this very rapid stream of articulatory gestures (which is speech-specific) enables enormous amounts of information to be compressed by a speaker into a single breath group, and to be then held in the listener's working memory, while its syllabic structure and meaning are decoded.

The Acoustic Theory of Speech Production [21] proposes that sound production in speech arises from the excitation of an acoustic filter composed of the

VT (pharynx, oral and nasal cavities) by a sound source (for example, the vibrations of the larynx for vowels). The source and filter are considered as linearly independent and the filter shape may be varied by motion of the articulators (tongue, teeth, lips and velum) to vary its acoustic properties. For vowels, different articulator configurations give rise to a VT filter with different resonances, known as formants, and hence to different phonemes. The formants, in particular the two with the lowest frequencies, F1 and F2, are critical to the identification of the perceived vowel with one or another target vowel category (/a/, /i/, /u/, etc.). The space delimited by the range of F1 and F2 values an individual or species can achieve is known as their *vowel space*. In MHs, this space is approximately triangular and the corners are associated with the *quantal* vowels: /a/, /i/ and /u/ [22]. To a first approximation, the acoustic properties of the VT for a given target vowel may be represented by two concatenated tubes of different lengths and cross sections and the frequency of each of F1 and F2 may be estimated, at least for the quantal vowels, from the geometric properties of one of these pairs of tubes [23].

The quantal vowels /a/, /i/ and /u/ are distinctive in that their acoustic pattern is perceptually stable across a reasonably large range of variability in the precise points of articulation, and perhaps for this reason, these tend to be the most common of the targeted vowel sounds used in human languages. Lieberman argues (from the acoustic analysis of simulated VT resonances) that production of the quantal vowels requires a VT anatomy that enables independent constriction of two cavities, front and back, and which can produce abrupt transitions between open and constricted sections at or near the midpoint of its length (with a ratio of cross-sectional area of the order of 10 : 1). He points out that the human adult VT is uniquely well-adapted for this, because the lowered larynx and hyoid bring the posterior tongue down into the pharynx, so that this part of the tongue dorsum can act as a movable anterior pharyngeal wall (independently constricting or enlarging the back cavity, the latter by contracting genioglossus and related muscles). The right-angled bend in the tongue dorsum at the back of the oral cavity separately enables constriction at that point (by contracting styloglossus to approximate the tongue surface towards the nasopharynx). The front cavity can meanwhile be independently constricted or enlarged by movement of the tongue blade and by opening or closing the jaw. Lieberman argues that non-human primates (monkeys and apes) cannot produce these quantal vowel sounds because their tongues are horizontally oriented and located entirely in the oral cavity. This anatomical pattern limits phonetic potential to a single-tube model: the tongue's intrinsic muscles and elastic properties then mean that other primates cannot achieve sufficiently sharp discontinuities in cross-sectional area near the midpoints of their VTs to produce human-like 'two-tube' vowel formants.

Lieberman proposes that the distinctive descended larynx position in the MH VT must be an adaptation to speech, because no other selective advantage for a

lowered larynx could have outweighed the disadvantage of a permanent separation of epiglottis from velum (which increases the risk of accidentally choking when a swallowed object gets lodged in the pharynx—Lieberman was composing his theory at a time when the Heimlich manoeuvre was very much in the news). Lieberman & Crelin [24] undertook an analysis of the reconstructed VT anatomy of Neanderthals. The Neanderthal larynx was placed high in the neck by analogy with the configuration found in non-human adult primates and newborn humans, on the basis that numerous aspects of the Neanderthal cranial base and mandible were more like the human newborn than adult forms.

Modelling of the potential acoustic capability of Neanderthal VTs was carried out by Lieberman *et al.* [24,25] based on their three-dimensional reconstruction of the La Chapelle-aux-Saints fossil. Silicon rubber casts of the VT of specimens of an adult male and a newborn MH and of a chimpanzee were made together with a cast of the predicted VT of the Neanderthal specimen as reconstructed by Lieberman & Crelin [24]. They noted that, owing to the assumption that the Neanderthal larynx is placed relatively higher in the VT than that of an adult MH, the Neanderthal tract is more similar to that of the newborn MH and chimpanzee than to that of the adult MH specimen. Acoustic modelling was confined to predicting the first three formants of the quantal vowels, /a/, /i/ and /u/. Area functions (the cross-sectional area at each point along the length of the VT) for the configurations of the chimpanzee VT said to be the 'best' approximations to /a/, /i/, /u/ were derived, and the corresponding formant frequencies were estimated using the algorithm developed by Henke [26]. Similarly, the formants for the newborn MH and the Neanderthal specimen were estimated. In each case, a comparison was made between the estimated formants and those measured by Peterson & Barney [27] for the American English vowel set in 76 children and adults. It was observed that the chimpanzee and newborn MH VTs both predicted substantially reduced vowel spaces compared with the Peterson and Barney data. The prediction for the Neanderthal vowel space was also smaller than for the MH dataset, being closer to that for the chimpanzee and newborn human specimens.

Lieberman & Crelin [28] concluded that Neanderthals could not produce the quantal vowel forms (/a/, /i/ and /u/) because they lacked independent variability of the pharyngeal cavity (which humans achieve by antero-posterior movements of the dorsal tongue). Lieberman [29] listed other fossil skulls that were morphologically similar to either the La Chapelle-aux-Saints Neanderthal or to the MH configuration, indicating that in each case the phonetic potential was expected to match that of the relevant comparator.[1] As more direct supporting evidence for his reconstruction, and having previously proposed that the Neanderthal styloid process (as also in human newborns) is characteristically less vertically aligned than in adult humans reflecting the more superior position of the point of insertion of the Neanderthal stylohyoid muscle, in a subsequent paper, Lieberman [29] also proposed that if the mandibular facets at the origin of

the anterior digastric, insertion of the posterior digastric, and origin of the geniohyoid muscles are angled to minimize sheer stress, then the human adult chin could be seen as an adaptation to the lowered position of the hyoid (since it enables these facets to be more inferiorly oriented); in Neanderthals as in human newborns, however, these facet orientations are consistent with a horizontal alignment of these muscles—again implying a more superiorly positioned hyoid. Lieberman, therefore, argued that selection on VT anatomy for stable and complex speech output had occurred among early MHs, but not among Neanderthals.

A number of studies have been carried out to test particular aspects of Lieberman's hypothesis, of which we summarize here only those most relevant to the focus of the remainder of this paper.

The argument that it is the resting profile of the tongue and the position of its muscle attachment points, rather than its fundamental muscle architecture, that differentiates the articulatory potential of the human and chimpanzee tongues has been supported by Takemoto [30,31], who has shown by dissection that the basic organization and orientation of muscle fibres is the same in both species. The major difference between the human and the chimpanzee tongue is the greater antero-posterior curvature of the surface of the human tongue, which gives it more degrees of freedom in feeding and in articulatory manoeuvres: the chimpanzee tongue, being essentially flat along most of its length, is largely restricted to protrusive and retrusive movements. However, the argument that the resting profile of the MH tongue must be explained by an evolutionary descent of the larynx under selection for speech capability has been separately addressed in a series of studies of the developing monkey, chimpanzee and human VT by Nishimura *et al.* [32–36], and these new studies have tended not to support that hypothesis. Nishimura *et al.*'s studies suggest that the two-tube configuration identified by Lieberman as central to the human VT's phonetic potential may have evolved as a secondary consequence of changes in mandible shape and of facial flattening, and not as a primary object of natural selection for phonetic potential.

With regard to Neanderthal/MH contrasts, initially it had been suggested that basicranial flexion (which is greater in MHs than in Neanderthals) might also be a marker of laryngeal descent [24,37,38], but subsequent anatomical studies have disproved this [39,40]. A different approach to Neanderthal VT reconstruction was taken by Boë *et al.* [41] who maintained that the Neanderthal larynx and hyoid bone sat lower in the VT than in Lieberman's reconstruction (using a prediction model for larynx height based on its correlations with various skull and mandible dimensions in an MH reference sample). Implications for speech were examined using a statistical approach to generating the maximal Neanderthal vowel space, based on the articulatory model of Maeda [42]. This method defines the principal components that underlie a defined proportion of the variance in a set of observed articulatory gestures in a corpus. By systematically covarying these components within a defined range, it is possible to predict all possible

articulatory gestures, within that range. Boë *et al.* [41,43] have modified Maeda's model to allow for adjustment of the laryngeal height for any particular VT by means of a laryngeal height index: broadly a coefficient relating (horizontal) palatal distance to (vertical) pharyngeal height. They used an acoustical model [44] based on the set of possible VT geometries to explore the maximal F1–F2 vowel space for MHs ranging in age from newborn to adult (male and female) as well as for an adult Neanderthal VT geometry estimated from the La Chapelle-aux-Saints fossil as reconstructed by Heim [45]. They estimated the larynx height index of the Neanderthal to be of the same order as that of a 10 year old child (in other words, with a long oral cavity relative to laryngeal height, the Neanderthal oral cavity being some 2 cm longer than in the reference adult human). This value for the index does not prevent 10 year old children from producing quantal vowels, and Boë *et al.* (having no reason to assume that it would have prevented Neanderthals from articulating quantal vowels) concluded that the modelled Neanderthal maximal vowel space did not significantly differ from that of an adult MH.

De Boer & Fitch [46] have critiqued Boë *et al.*'s approach on the grounds that their focus on estimating Neanderthal laryngeal height does not properly validate the inference of a similarly lowered position of the hyoid and tongue root. Further they suggest that, because Boë *et al.* have started from a theoretical model [47] that incorporates all possible MH VT shapes, the estimates of the Neanderthal vowel space they obtain from it may be biased towards a MH-like vowel triangle. Meanwhile, de Boer [48,49] has investigated the effect of larynx height on potential vowel space using Mermelstein's [50] model of the direct motion of the articulators, which is based on the integrated effect of the action of their associated muscles. This model permits direct control of the location and/or shape of the larynx, pharynx, hyoid, velum, tongue body, tongue blade, jaw and lips. De Boer considered VT mid-sagittal cross sections corresponding to an adult male MH and an adult female MH (and also a 'mixed' male VT model with female perilaryngeal anatomy at the level of the larynx). Area functions were generated from the mid-sagittal cross sections using the formulae given by Mermelstein [50] derived from X-ray data. A large number of VT configurations (10 000) were generated by random selection of the control parameters of the model to define the potential vowel space. Where a parameter set resulted in occlusion of the VT, the area was re-set to 0.1 cm^2. The formant frequencies predicted using Kelly & Lochbaum's [51] method were plotted on an F1–F2 diagram and an estimate of the complete vowel space was made by calculating the convex hull of the cloud of data points so derived. The results showed that the female VT was predicted to have a larger vowel space than the male VT, given the same articulatory constraints. De Boer concluded that a VT with approximately equal lengths for its horizontal and vertical segments was optimal for maximizing articulatory range, and that the further descent of the larynx in the MH post-pubertal male must have been driven by a selection pressure other than enhancement of speech communication (such as size exaggeration).

In this paper, we report a new study of the speech potential of Neanderthals, based on reconstructions of the positions of the hyoid and tongue root (and thus addressing the critique of Boë *et al.* [41] made by de Boer & Fitch [46]), and a software articulatory model that also allows us to explore the effects on vowel space of varying the position of the hyoid and other articulators. Data obtained in other domains (e.g. handedness as inferred from tools and skeletal parts; ancient DNA) provide suggestive but indirect evidence of the potential for complex vocal communication in Neanderthals. We propose our new approach as a way of incrementally building understanding of more direct evidence for Neanderthal speech potential, while avoiding the polarization and polemic that have tended to characterize this debate.

2. MATERIAL AND METHODS
(a) *Anatomical reconstructions*
For the Neanderthal VT reconstructions, we predicted cranio-caudal and dorsoventral distances to a hyoid landmark from skull and mandible landmarks or reference planes. These Neanderthal hyoid position reconstructions are based on prediction equations obtained by analysing a human reference sample, which use three-dimensional interlandmark distances describing face, skull base and mandible dimensions as the independent variables.

(i) *Human sample*
Our MH data consist of computed tomography (CT) scans produced from 10 female and 10 male volunteering healthy adults aged between 20 and 65. The data were provided by the Laboratory of Functional Anatomy of the University Paris René Descartes [52]. Study subjects were scanned in dorsal decubitus position, teeth in occlusion and tongue held against the palate. Image acquisition occurred parallel to the Frankfurt plane. CT scans were acquired using a Somatom Sensroation16—Siemens scanner (slice thickness 0.75 mm, pixel matrix 512 × 512, Inc. 0.48 mm, tube voltage 120 kV, tube current base line 200 mA, FOV 160 mm) at Ouest Parisien Medical Imaging Centre. Based on the exploration of the range of dorsoventral and cranio-caudal hyoid positions observed in this MH sample, we defined an envelope of anatomically observed human vertical hyoid positions in relation to the vertebral column, and horizontal hyoid positions in relation to the skull base (see electronic supplementary material, figure S1). Neanderthal hyoid position reconstructions were considered more anatomically plausible if they fell within the boundaries set by the MH sample projected onto the Neanderthal skulls and vertebral column reconstructions (see below).

(ii) *Neanderthal sample*
The fossil sample consisted of CT scans (industrial and medical scanners, various sources) and surface laser scans (NextEngine 3D Scanner, accuracy 0.4– 0.7 mm) of skulls and/or mandibles of eight adult (La Ferrassie, La Chapelle-aux-Saints, Gibraltar 1, Guattari, Shanidar 1, Abri Bourgeois, La Quina 9 and Regourdou) and one subadult (Le Moustier) specimen

and one case of an adult hyoid (Kebara). CT scans and stereolithography interface format (STL) models were acquired from museum casts and reconstructions or—where possible—from original bones. In cases where the fossil specimens had not already been fully reconstructed and/or crucial pieces were missing, we conducted additional reconstructions. This was necessary in order to obtain the full set of measurements. Two types of reconstructions were applied. We added missing pieces using boundaries defined by the contour of the remaining bone material. This technique was useful to reconstruct a lost coronoid process or mandibular angle (gonion) or nasal bone (rhinion). For specimens consisting of partial mandibles or skulls, we used mirror image techniques to copy the surviving pieces. The mirrored pieces were then fitted together in the three-dimensional image space using anatomical expertise and a best-fit approach.

(iii) *Computed tomography scan data processing, distance measurement definition and collection*

From the CT scans, we segmented three-dimensional representations of the skull, mandible and—in the case of the MH reference sample—the hyoid bone. From the STL models, we simply produced three-dimensional surface representations. The software suite AMIRA (Visage Imaging) was used in both cases. We chose 35 standard osteological measurements describing face, cranial base and mandibular dimensions (see electronic supplementary material, table S1). On the MH sample, we also defined and collected 19 hyoid distance measurements (12 vertical and seven horizontal, see table 1), which describe the position of the hyoid in relation to landmarks on the mandible and skull, measured parallel or orthogonal to the Frankfurt plane. The reference point on the hyoid was always the posterior-most point in the mid-sagittal plane on the superior rim of the hyoid body. All measurements (both mandible and cranial dimensions and hyoid distance measurements) were collected after skulls were orientated in the Frankfurt plane, using the Software suite AMIRA (Visage Imaging) and TPSdig2 [53] for measurement collection.

(iv) *Human regression models and application to Neanderthal reconstructions*

Neanderthal hyoid positions were reconstructed using prediction equations (multiple linear regression models) obtained in an analysis of the MH sample. From an original set of 19 such prediction equations we retained five to use in predicting Neanderthal hyoid position (i.e. the offset distances to a mid-sagittal hyoid landmark from a skull or mandible landmark/reference plane in the x and y axes, dorsoventral and cranio-caudal). Selection criteria for these five regression models were as follows: R^2 values greater than 80 per cent, and all skull and mandible dimension measurements retained (following stepwise variable selection) as independent variables in the regression equations had to be based on landmarks which survived and could be identified on the three-dimensional scans of the fossil skulls.

We then used the predicted hyoid distances for two adult Neanderthal specimens (La Ferrassie and La Chapelle-aux-Saints (La Chapelle)) to reconstruct

Table 1. Coefficients of determination (R^2), expressed as a percentage, for all hyoid distance regression models in the adult human reference sample. SS, based on stepwise selection from skull-only set of possible independent variables; FS, based on stepwise variable selection from full set including mandible dimensions; bold type, models selected for further analysis according to best-fit criterion, subject to the requirement that the dependent and independent variables also reference landmarks, which survive and can be identified on the three-dimensional scans of the fossil skulls.

orientation	area	distance	human R^2
vertical	mandible	hyoid–gonion	(FS) 78.3%
		hyoid–condyle	**(FS) 80.1%**
		hyoid–coronion	(FS) 79.8%
		hyoid–infradentale	(FS) 97.4%
		hyoid–gnathion	(FS) 88.5%
	skull	**hyoid–palatal plane**	**(FS) 64.9%**
		hyoid–porion	**(FS) 89.6%**
		hyoid–sella	(FS) 80.3%
		hyoid–nasion	(FS) 88.5%
		hyoid–prosthion	(FS) 96.9%
		hyoid–ANS	(FS) 96.3%
		hyoid–rhinion	**(FS) 92.6%**
horizontal	mandible	hyoid–gnathion	(FS) 75.1%
		hyoid–infradentale	(FS) 69.5%
	skull	hyoid–nasion	(FS) 41.9%
		hyoid–prosthion	**(SS/FS) 91.5%**
		hyoid–basion	(FS) 66.0%
		hyoid–PNS	(SS/FS) 58.0%
		hyoid–ANS	(FS) 89.3%

and visualize the Neanderthal three-dimensional hyoid position. This also served as a test to see whether predicted Neanderthal hyoid positions fell within the position envelopes observed in the MH sample in relation to skull and vertebral column. We did not have access to CT scans of Neanderthal cervical vertebrae; however, earlier studies show that humans and Neanderthals are very similar in overall cervical spine length [54,55]. Therefore, the largest male cervical spine including C1 to C6 from our *in vivo* human sample, was selected and used as a proxy for the Neanderthal visualizations. In the three-dimensional shape space, we placed the human cervical spine in the correct anatomical position and in contact with the cranial condyles under the Neanderthal skulls orientated in the Frankfurt plane. For the Neanderthal hyoid, we used an STL scan of the Kebara hyoid [10].

(b) *Acoustic modelling*

Simus_Neanderthals is a software modelling tool that allows evaluation of potential Neanderthal VT geometries in terms of estimates of the vowel space they can produce. It draws on previous data and software developed for modelling human speech articulation by Badin and co-workers (cf. [56]). The model is flexible, allowing a variety of hypotheses to be tested depending on the selected modelling parameters and assumptions. The output of the model is an estimate of the formant frequencies F1 and F2 of a vowel from a Neanderthal VT analogous to that specified by a given reference
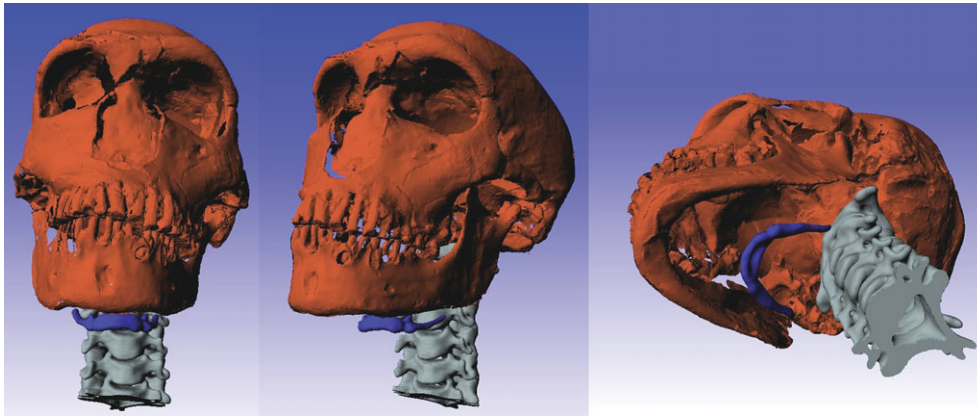
Figure 1. Three views of the Neanderthal mesh based on high-resolution CT scans of the fossil sample La Ferrassie. The hyoid bone mesh comes from an STL scan of the Kebara hyoid [10] and its location is based on the human reference equation as described in §3. The mesh of cervical spine is a scan of a human sample located within the Neanderthal mesh using anatomical landmarks as described in §2.
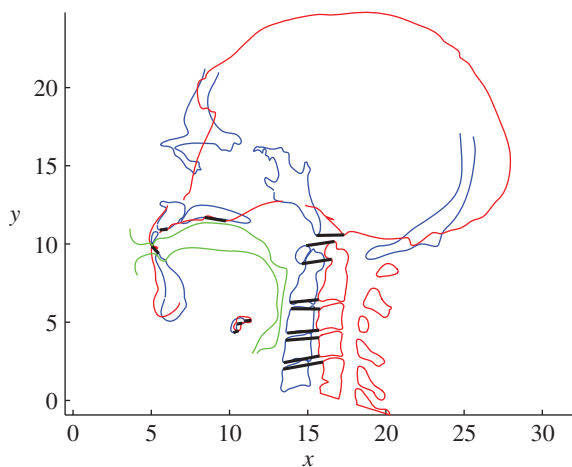

Figure 2. Mid-sagittal cross section through the Neanderthal mesh (red) and mid-sagittal cross section of corresponding modern human bony architecture (blue) with its associated modern human vocal tract (green). Black lines show the location of the 16 selected bony landmarks in each sample from which the transformation vectors are derived.

human VT. Modelling starts from the three-dimensional mesh models of Neanderthal bony anatomy of the skull, jaw and hyoid together with a human cervical spine sample as described above and shown in figure 1. The use of the human spine is discussed above.

The location of the hyoid bone for the Neanderthal mesh and the degree of opening of the jaw are user-selectable parameters of the model, and choosing different positions for these features allows testing of hypotheses about the effect of varying their location on the acoustical output from the VT. To make the acoustical model, the mesh is first cut in the mid-sagittal plane. The resulting outline is plotted on an arbitrary reference grid with the lower edge of the upper incisors located at $(x = 5, y = 10)$. The x-axis is oriented from anterior to posterior and approximately in the occlusal plane, and the y-axis from feet to head. A mid-sagittal section of a reference human skull, hyoid, jaw and spine is plotted on the same reference grid with the same orientation and incisor location as shown in figure 2. A reference human VT is associated with the reference human bony architecture.

For the Neanderthal (N) and the reference human (R) cross sections, 16 landmarks are identified in the bony architecture to form a set of transformation vectors. Selection of the landmarks was based on the set of most clearly identifiable features within the bony architecture of the Neanderthal and the reference human. The set of landmarks consists of three points on the hard palate (one at the most anterior point, one at the most posterior and one midway in between), eight points on the anterior upper and lower corners of each cervical vertebra, one point on the lower margin of the skull (occipital condyle), three points on the hyoid bone (the most inferior point, the most posterior point and the centre of the area) and one landmark on the most superior point of the anterior mandible. The user may choose whether the landmarks are considered in groups or individually according to the preferred modelling assumptions. A spatial transformation vector from each R landmark to the corresponding N landmark is calculated. If landmarks have been grouped, the mean of the transformations for each landmark in the group is applied to all members of the group.

Next, the reference VT is divided into user-defined sections, and each section is associated with a landmark or group of landmarks. The decision regarding which section of the VT is associated with which landmark or group of landmarks forms a further set of modelling assumptions. Each section of the VT is then deformed, point by point, using the transformation for its associated landmark or landmark group. Following transformation, junctions between VT sections are linearly smoothed to eliminate any boundary discontinuities or abrupt geometry changes. A boundary condition preventing the tongue crossing the upper VT margin is also applied. These two processes are jointly referred to as *boundary conditioning*.

From the transformed VT, an area function is calculated by sectioning the VT into a sequence of short, contiguous, cylindrical tubes using a standard grid (e.g. [21,44]) and then applying the alpha-beta model of Soquet *et al.* ([57], p. 176, table 3) for an adult male to obtain estimates for the cross-sectional area of each tube based on the height of its sagittal section. Any completely occluded sections of the tract are
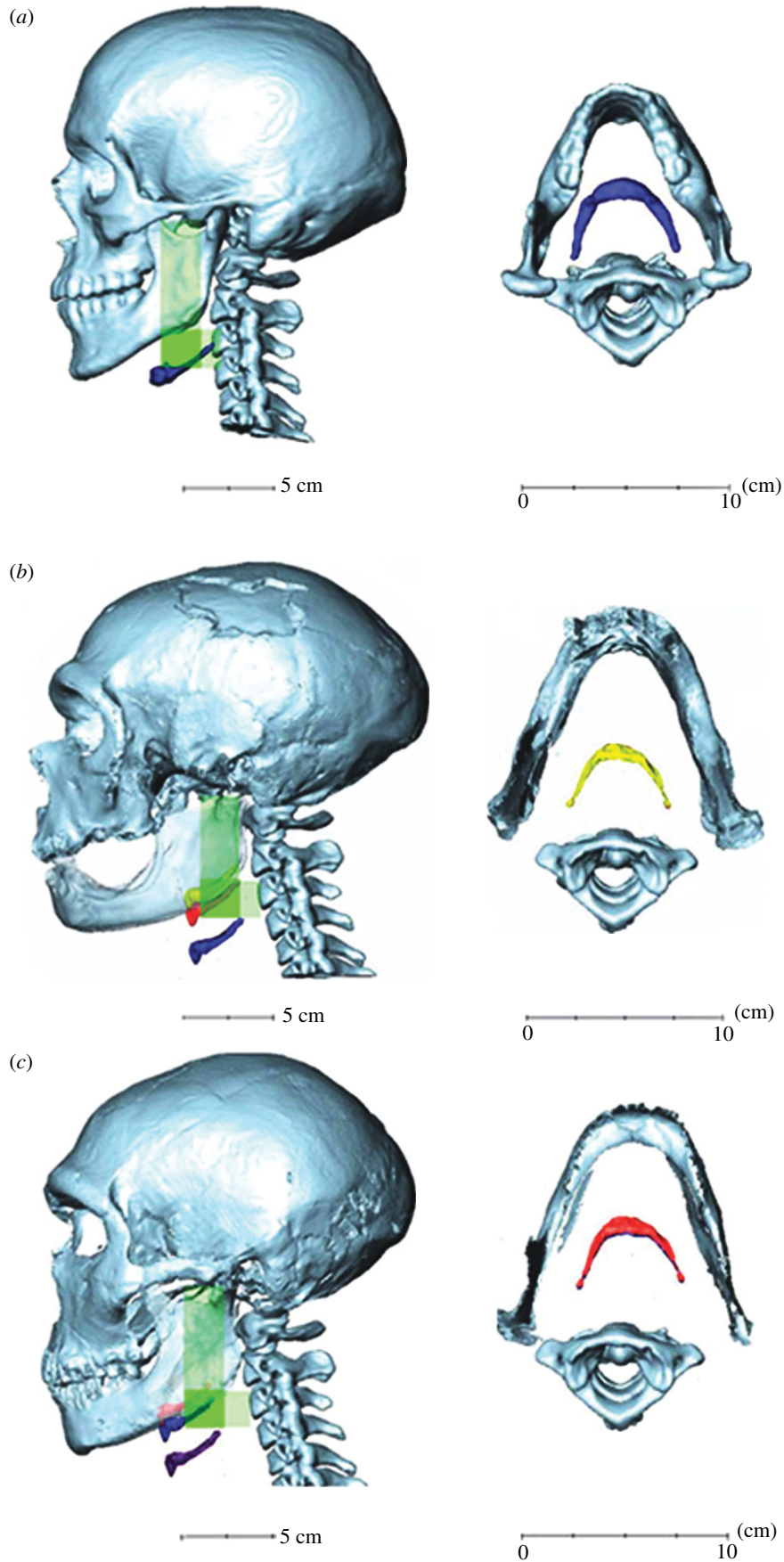
Figure 3. (*a*) Hyoid position as observed in the MH adult reference sample, and (*b,c*) as predicted for two adult Neanderthal specimens ((*b*) La Chapelle-aux-Saints; (*c*) La Ferrassie). The green shading shows the range of positions of the hyoid observed in the adult human reference sample, measured as vertical and horizontal offsets from skull base and cervical vertebral land-marks. The coloured images of the hyoid in the Neanderthal specimens are positioned as predicted from alternative human reference equations (tables 1 and 2; hyoid–condyle, blue; hyoid–porion, red; hyoid–palatal plane, yellow; hyoid–rhinion, purple). The predicted vertical offsets are shown in profile view in the left-hand column, and the predicted horizontal offset (estimated from the single-used regression model) is illustrated in plan view in the right-hand column.

then reset to a minimum area of 0.2 cm². The output of this stage of the modelling process is the estimated N area function derived from the transformed R VT geometry. Linear acoustic modelling (e.g. [58]) is then applied to the area function giving an estimate of the formant frequencies for the Neanderthal analogue under the chosen set of modelling parameters and assumptions.

3. RESULTS

(a) *Results 1: Neanderthal vocal tract reconstruction*

(i) *Modal qualitative horizontal and vertical hyoid positions in human computed tomography reference sample*

The empirically observed range of the adult human hyoid positions in our CT reference sample was recorded in relation to landmarks on the cranial base (horizontal offset) and cervical spine (vertical offset; figure 3). In the antero-posterior axis, the modal human position for the reference point on the hyoid body is found in a plane passing through the posterior nasal spine orthogonal to the Frankfurt horizontal (55% of the sample) or slightly (less than 5 mm) anterior or posterior to it (35%). This plane also bisects the mandible at the deepest point of the intercondylar notch. In the supero-inferior axis, the reference point on the hyoid body is mostly found aligned on a plane parallel to the Frankfurt horizontal that passes through the inferior endplate of cervical vertebra C3 (females) or the superior endplate of C4 (males).

(ii) *Reconstructed Neanderthal hyoid position*

The human regression models (tables 1 and 2) resulted in Neanderthal estimates with similar patterns of hyoid to skull and mandible distances as those observed in the human sample (table 3 and figure 3). We have also included vertical distances to hyoid from the palatal plane, although the fit of the regression model is quite poor in the human reference sample, because this is a plane that has been used for similar purposes in previous studies [41]. The predicted means for Neanderthal vertical hyoid distances were not significantly different to the observed mean distances in the adult human sample, but the predicted mean distances to the hyoid in the antero-posterior axis were significantly greater in Neanderthals (table 3).

For the subsequent acoustic analyses, we chose the most MH-like hyoid position as reconstructed using the hyoid–porion (cranio-caudal) and the hyoid-prosthion (dorsoventral) regression equations. The hyoid–porion model was chosen because it resulted in both the La Ferrassie and the La Chapelle-aux-Saints specimen in a cranio-caudal hyoid position which fell within the cranio-caudal hyoid position range in relation to the vertebral column observed in the MH sample, when this was projected onto the mid-sagittal plane in those two Neanderthal digital three-dimensional images of the skull and mandible (with the human vertebral column added as described in §2; see figure 3).

(iii) *Visualization of reconstructed Neanderthal hyoid position*

In La Chapelle, three-dimensional reconstruction of horizontal hyoid position (figure 3) based on the human reference equations places it within or very close to the range found in the human reference sample. Vertical distances fall within the distribution space of the human reference sample with the exception of hyoid–condyle and hyoid–rhinion. This reflects the weight given by those prediction equations to this fossil's very wide mandible (GOW) and long cranial base (FL2). The predicted horizontal hyoid distance of the La Ferrassie Neanderthal three-dimensional reconstruction (figure 3) based on the human reference equation is slightly anterior to the positions recorded in the human reference sample, reflecting this fossil skull's relatively long cranial base and oral cavity (which have negative weights in the reference human multiple regression equation). Vertically, the predicted hyoid distance is within the range for the human reference sample with the exception of hyoid–rhinion distance, which has an excessively high predicted value. This is because cranial dimensions in the regression model include total face height (THF), piriform aperture height (PAH) and upper face height (UFH), all of which contribute substantially to facial shape differences between humans and Neanderthals.

(b) *Results 2: acoustic analysis*

As an example of how the model may be used, we can consider a specific case study using the Neanderthal mesh from the La Ferrassie sample (shown previously in figures 1 and 3) and a MH reference. Three VT reference configurations will be considered, one for each of the quantal vowels /a/, /i/ and /u/ [22], which are considered to acoustically delimit the extremes of the attainable MH vowel space [59]. The output of the model is an estimate of the formants of the analogue of these vowels for the Neanderthal VT. The effect on the Neanderthal analogue of these vowels owing to varying the jaw opening and the hyoid position will be explored. Initially, the hyoid bone for the Neanderthal VT was placed in the location designated anatomically predicted (i.e. the position reconstructed using the hyoid–porion (cranio-caudal) and the hyoid–prosthion (dorsoventral) regression equations.). The 16 bony landmarks were grouped into six groups designated as *palate tip* (one landmark), *palate mid* (one landmark), *palate back* (one landmark), *spine* (nine landmarks), *hyoid* (three landmarks), and *jaw* (one landmark), respectively. The VT sections were associated with landmark groups as shown in figure 4. This figure represents the raw transformation prior to boundary smoothing and conditioning. The transformed mid-sagittal VT section for each of the vowels after boundary conditioning is shown in figure 5 together with the MH version used to obtain the transform.

For each transformed VT, the area function was then calculated as described in the previous section. Examples are shown in figure 6a together with the corresponding MH area functions. Also calculated and shown in figure 6b are the acoustic transfer function estimates for the Neanderthal and MH VTs. These indicate the energy that would be transferred from laryngeal vibrations to the lips for each frequency and the peaks correspond approximately to the formant frequencies of the tract.

Table 2. Summary of multiple regression equations for the five hyoid position prediction models selected for further analysis. All five models were statistically significant at $\alpha = 0.001$. Abbreviations: MFD, middle face depth; GOW, gonion width; OCWmd, oral cavity width (mandibular); ZMW, midface width; ANS–hormion, anterior nasal spine to hormion; THF, total face height; PAH, piriform aperture height; UFH, upper face height; FML, foramen magnum length; FL2, face length2; IOW, interorbital width; OFD, total face depth; PAW, piriform aperture width; OCL, oral cavity length; PHL1, oropharynx length1.

hyoid distance	regression equation
condyle	$= (1.399 \times \text{MFD}) + (0.422 \times \text{GOW}) + (-6.512)$
porion	$= (0.884 \times \text{MFD}) + (0.539 \times \text{GOW}) + (-0.412 \times \text{OCWmd}) + (0.318 \times \text{ZMW}) + (-4.259)$
palatal plane	$= (1.366 \times \text{ANS–hormion}) + (-0.81 \times \text{OCWMd}) + (-0.2)$
rhinion	$= (1.035 \times \text{THF}) + (2.23 \times \text{PAH}) + (-2.16 \times \text{UFH}) + (1.055 \times \text{FML}) + (-2.95)$
prosthion	$= (1.726 \times \text{FL2}) + (1.427 \times \text{IOW}) + (-1.519 \times \text{OFD}) + (1.658 \times \text{PAW})$ $+ (-0.431 \times \text{OCL}) + (0.326 \times \text{PHL1}) + (1.087)$

Table 3. Summary of mean hyoid distances (in centimetres) between humans and Neanderthals. Significant differences $p < 0.01$ (i.e. 0.002, Bonferroni corrected) shown in italics; no other differences exceeded the Bonferroni-corrected $p < 0.05$ threshold. Mean observed human distances based on $n = 20$; mean predicted Neanderthal distances obtained from human regression models based on $n = 3$ for all distances except for hyoid–prosthion distance, $n = 6$. PP, palatal plane.

mean distance	condyle (vertical)	porion (vertical)	PP (vertical)	rhinion (vertical)	prosthion (horizontal)
humans	7.058	7.972	5.617	9.845	*6.724*
Neanderthal	7.81	8.081	4.516	10.643	*8.296*
difference	0.752	0.109	−1.101	0.795	*1.572*

In figure 7, the red triangle shows the vowel space for the reference VT (MH) and the dark blue triangle shows the vowel space for the VT after transformation (N). The axes are the values of the formants F1 and F2 and are plotted on the Bark scale which ranges from 1 to 24 Barks, corresponding to the first 24 critical bands of MH hearing with the conversion from a frequency, f, to the equivalent value in Bark, b, given by:

$$b = 13 \arctan(0.00076 \times f)$$
$$+ \ 3.5 \arctan\left(\left(\frac{f}{7500}\right)^2\right). \tag{3.1}$$

Note that implicit in this representation is an assumption that it is reasonable to plot the predicted Neanderthal formants on the Bark scale, which is based on the psychoacoustic perceptual capabilities of the MH ear. It is of course possible that a Neanderthal hearing system would have a different psychoacoustic profile and the Neanderthal would perceive these vowels differently to a MH.

The ellipses (shown partially in the figure) are based on the Peterson & Barney [27] data for the vowels /a/, /i/ and /u/, which show the range of typical F1 and F2 values for a sample of speakers of American English. It is notable that for the transformed VT, /i/ and /u/ have formants close to the human reference values, but /a/ does not. Also plotted in the figure are the Neanderthal vowel triangle estimates from the predictions of Lieberman & Crelin [28] and Boë *et al.* [43].

As an example of the effect of changing the modelling assumptions, figure 8 shows the effect on the vowel triangle of moving the hyoid bone from the location designated *anatomically predicted* to locations with coordinates ± 0.5 cm in the x-direction and ± 1.0 cm in the y-direction relative to the *anatomically predicted* location.

Figure 9 shows the effect on the predicted vowel triangle of opening the jaw a distance of -0.5, -1.0 and -1.5 cm in the y-direction relative to the position in the original mesh. The hyoid bone was at the *anatomically predicted* location for each of these trials.

For illustrative purposes, in the electronic supplementary material, we include three sound files of the simulated adult human quantal vowels /a/, /i/ and /u/ and three sound files of the simulated Neanderthal quantal vowels /a/, /i/ and /u/ (obtained for the hyoid position reconstruction designated *anatomically predicted*; see blue triangle, figure 7).

4. DISCUSSION

Consistent with previous investigators [37,41], we have extrapolated soft tissue parameters of the Neanderthal VT from cranial and mandibular morphology, based on observations of their relationships in a human reference sample. Our results permit a reconstruction of the vertical position of the Neanderthal hyoid that places it within the range of empirically observed locations on this axis (the y-axis) in the human reference sample. However, our extrapolations of Neanderthal hyoid position in the antero-posterior axis (the x-axis) predict a distance to hyoid from the anterior margin of the oral cavity, which is significantly greater (by 1–2 cm) than in our human reference sample, and even then, the predicted location in relation to cranial base landmarks is further forward from the posterior pharyngeal wall than we would expect in a MH. This difference has also affected our acoustic analysis, since the
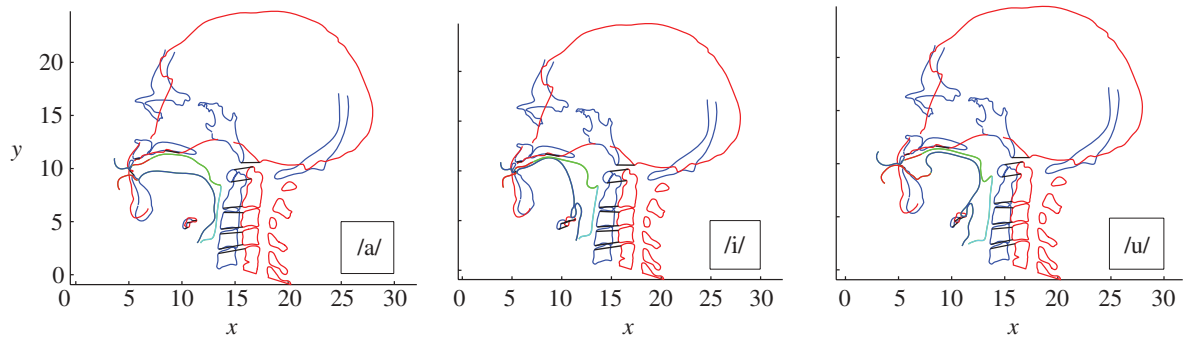
Figure 4. Association of VT sections for each quantal vowel with the bony landmark groups for the transformed VTs prior to boundary conditioning. Sections are associated with landmarks *tip of palate* (blue), *mid of palate* (red), *back of palate* (green), *spine* (cyan), *hyoid* (magenta) and *jaw* (violet), respectively.
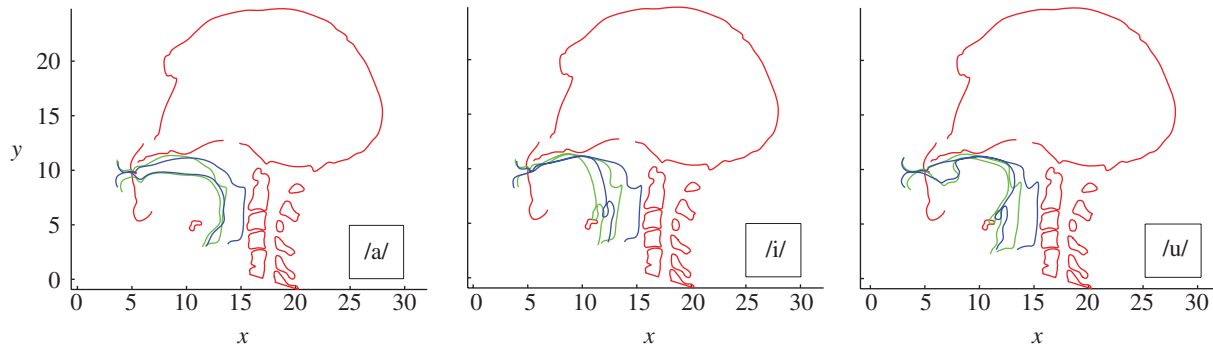

Figure 5. Transformed VT (blue) and original VT (green) for each of the quantal vowels /a/, /i/, and /u/ following boundary conditioning.

reconstructed anterior placement of the hyoid relative to the posterior pharyngeal wall limits constriction of the pharyngeal cavity in the /a/ vowel articulation. We do not believe that we have discovered a key physical property of the Neanderthal VT in this respect: rather, this difference reflects the limitations of our method of using a MH reference sample given underlying differences in the morphology of the Neanderthal skull and mandible. MHs are characterized by greater facial flattening (shorter oral cavities), although the reasons for this remain unclear [60,61]. This places limits on fine-grained inference of hyoid-craniofacial hard tissue spatial relationships in a fossil species, when our living reference model (MHs) differs in global skull shape properties that determine the soft tissue reconstruction.

For our case study using Simus_Neanderthals, with the jaw in its default position and the hyoid in the *anatomically predicted* location, figure 7 shows that F1 and F2 are close to the MH range for both /u/ and /i/. F2 is also close to the MH range for /a/, but for this phoneme, F1 is significantly under-predicted. One reason for this may be the relatively larger pharyngeal cavity predicted for the N VT when compared with the MH VT, observable in the top left of figure 6. For /a/ Apostol *et al.* [23] show an affiliation between the Helmholtz resonance of the back cavity, and F1 that would, therefore, tend to lower values for larger pharyngeal volumes. On the other hand, Badin *et al.* [62] show that /a/ is a focal point, where F1 affiliation can equally be a quarter wavelength resonance of the front cavity for /a/ and may switch

affiliation depending on the precise articulatory geometry adopted. It appears to be this affiliation to the front cavity that we are observing when, as the jaw opens, we see an increase in F1, consistent with the findings of Sundberg [63] for soprano singers. Figure 9 demonstrates this, although we were unable to find reasonable jaw opening widths that resulted in an F1 as high as that found in MHs. Further, in our case, as F1 increased with jaw opening, F2 increased above the Peterson & Barney [27] range for /a/.

Comparing our predictions with those of Lieberman & Crelin [28], neither our predictions nor those of Lieberman & Crelin produce a good match to a MH /a/ reference, but in our case, the value for F1 is significantly lower than expected, whereas their F1 compares well with the MH values. Their F2 prediction for /a/ gives a value higher than the MH range, whereas ours is within the ellipse. Comparing the area functions for /a/, the enlarged pharynx that we predict is not replicated in their model suggesting the basis for their acoustic prediction differs from ours, and perhaps has a different association between formants and cavities. The predictions of Boë *et al.* [43] are within the ellipses for all the quantal vowels. Clearly, the output of any given model is significantly dependent on the choice of modelling assumptions used to derive the VT geometry.

For our case study, figure 8 shows that moving the hyoid from a posterior to a more anterior position reduces F1 with the largest effect occurring for /i/, especially when the hyoid is also positioned low. Raising the hyoid also decreases F1, especially for /i/ when the hyoid is also positioned posteriorly. For F2,
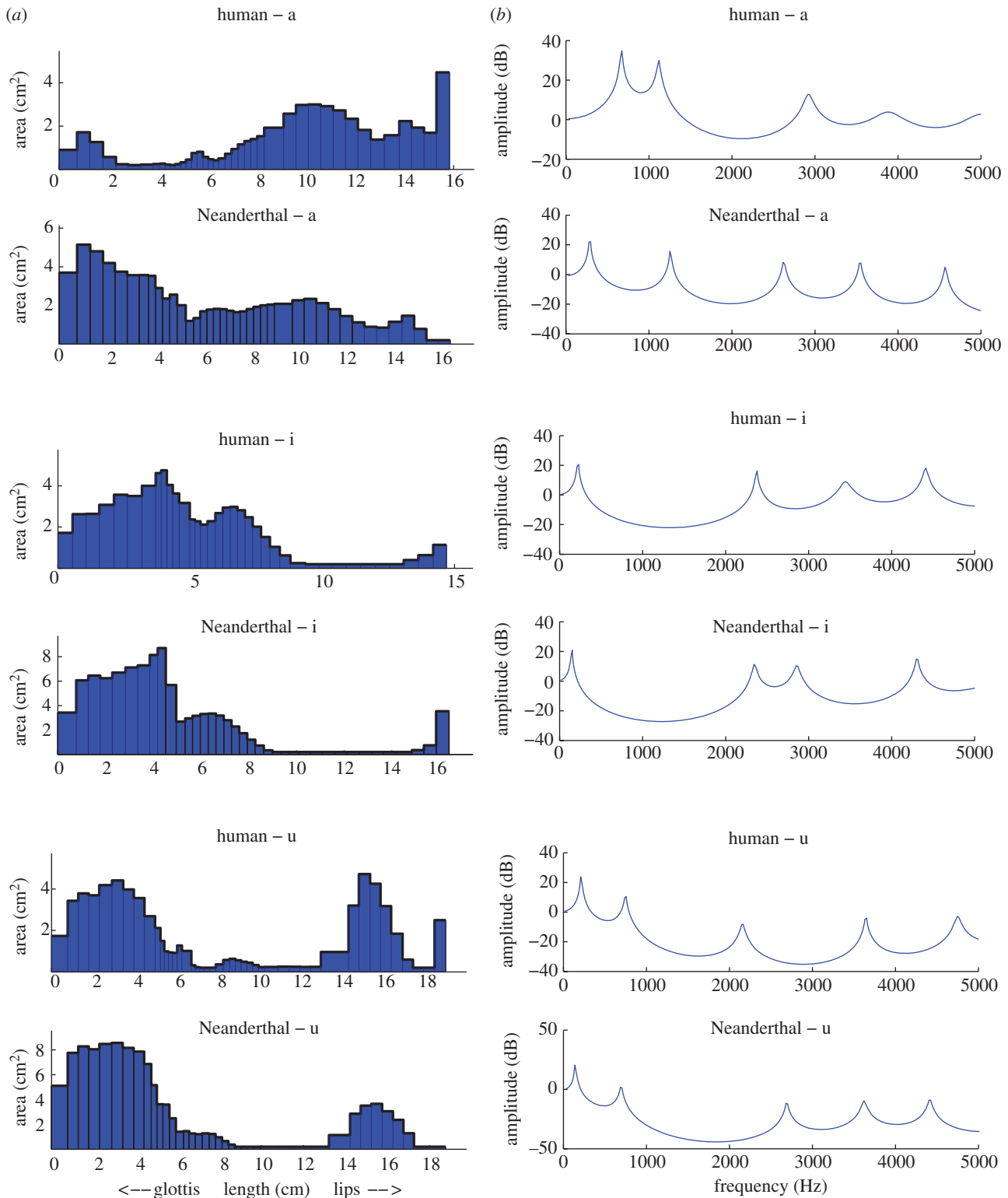
Figure 6. (*a*) Area functions for MH and transformed (Neanderthal) VTs and (*b*) corresponding VT acoustic transfer functions. The peaks in the transfer function correspond to the formant frequencies for a given VT area function.

moving the hyoid to a more anterior position causes an increase in frequency, with a larger effect for /a/ when the hyoid is also raised, and raising the hyoid causes a small increase in frequency. No reasonable position for the hyoid bone increases F1 for /a/ sufficiently to give a good comparison with the MH data under this set of modelling assumptions.

Opening the jaw for a fixed hyoid location (figure 9) has little effect on F1 for /i/ or /u/, but does increase F1 for /a/ quite significantly although it is still not raised to

MH frequencies. There is also an increase in F2 for both /u/ and /a/.

In figures 8 and 9, each predicted vowel triangle uses a fixed choice for hyoid location and jaw opening for all three phonemes. A best match between the vowel triangle for MH and N could perhaps be achieved by optimizing the pairing of jaw opening and hyoid location for each phoneme individually. Note that the predicted vowel triangle has corners that have the N analogue of the MH vowels. It may not be identical to the maximal
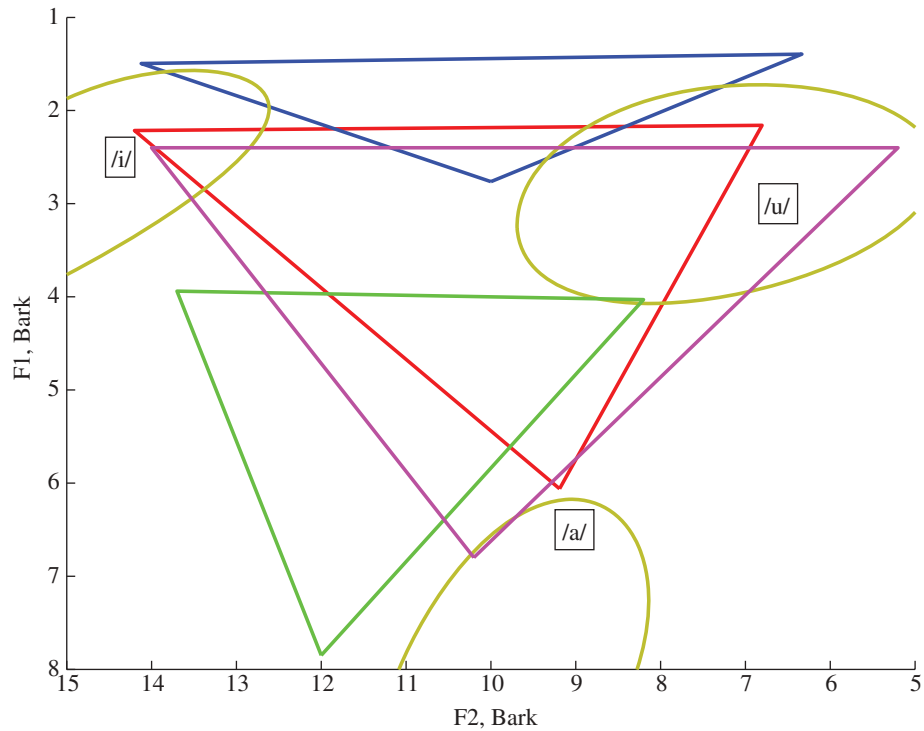
Figure 7. Vowel space for: the MH reference (red); the Simus_Neanderthals prediction (blue) with the hyoid bone in the *anatomically predicted* location and the default jaw opening; the predictions of Lieberman & Crelin (green [28]) and the predictions of Boë *et al.* (deep pink [43]). The gold ellipses (shown partially) represent the regions of the vowel space covered by each of the datasets for American MH quantal vowels recorded by Peterson & Barney [27].

Figure 8. Vowel space for: the MH reference (red); Simus_Neanderthals prediction with the hyoid bone in the *anatomically predicted* location (dark blue) and Simus_Neanderthals prediction with the hyoid bone displaced from the *anatomically predicted* location by: 1 cm superiorly and 0.5 cm anteriorly (pink), 1 cm superiorly and 0.5 cm posteriorly (mauve), 1 cm inferiorly and 0.5 cm posteriorly (grey-blue), 1 cm inferiorly and 0.5 cm anteriorly (pale blue). The jaw opening in each case is the default value and the gold ellipses (as in figure 7) represent the Peterson & Barney [27] data.

vowel space for a Neanderthal VT, since vowels with formants closer to the MH /a/, /i/ and /u/ may be achievable from different MH reference tracts.

When considering the results of this case study, it should be remembered that the hyoid location designated *anatomically predicted* in §3 is in fact only one of a
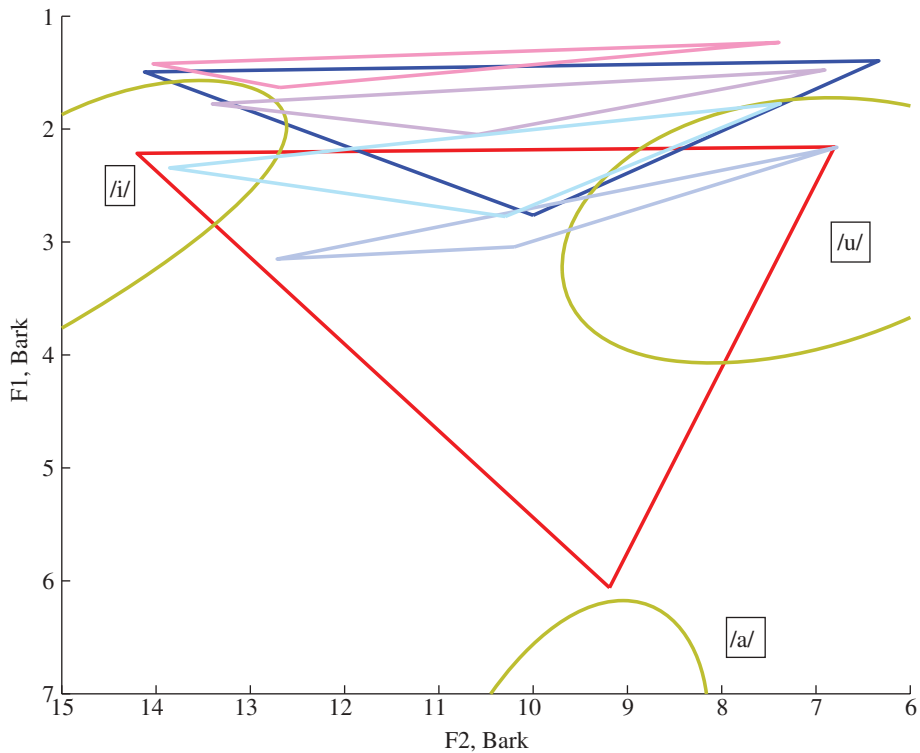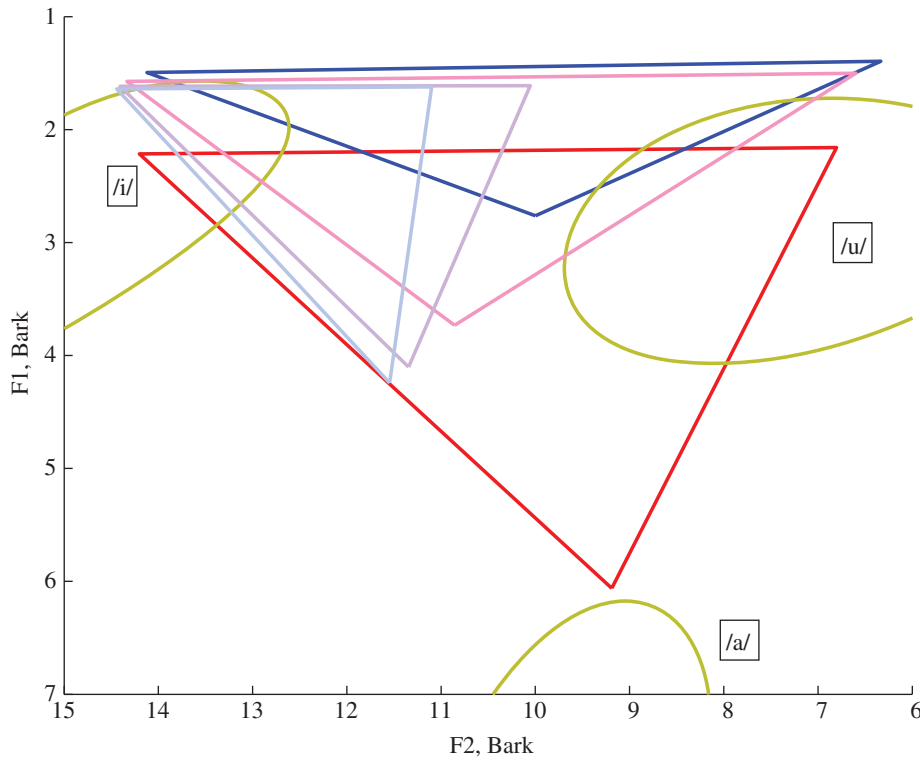
Figure 9. Vowel space for: the MH reference (red); Simus_Neanderthals prediction with the jaw in the default location (dark blue) and Simus_Neanderthals prediction with the jaw displaced inferiorly from the default location by: 0.5 cm (pink), 1 cm (mauve) and 1.5 cm (pale blue), respectively. The hyoid bone in each case was in the location designated *anatomically predicted* in §3 and the gold ellipses (as in figure 7) represent the Peterson & Barney [27] data.

range of possible locations predicted by the regression analyses. Although we have examined the effects of some horizontal and vertical displacement of the hyoid from that predicted by the chosen regression equation, we have not yet conducted a systematic sensitivity analysis of the acoustic and articulatory implications of positioning the hyoid at locations estimated from alternative regression models. Additionally, Simus_Neanderthals has a number of fixed parameters that may influence the outcome of the acoustical predictions. The 16 bony landmarks were chosen as identifiable locations observable in both the MH and N mid-sagittal cuts through the mesh. We have not been able to test whether this set is in any way either necessary or sufficient to achieve an optimum set of transformation vectors. We have chosen to use linear smoothing after VT transformation. No test has been made of smoothing algorithms other than linear, although the choice of smoothing method will affect the detail of the VT shape to some extent. Further, since anatomical considerations point to a smooth VT outline, it might be argued that the need to smooth at all is a facet of an inadequate transformation process. A similar case might be made regarding the hard constraint required to keep the tongue within the VT boundaries and the need to adjust for a minimum VT area when transformation predicts complete occlusion of the tract. Our treatment of these unrealistic predictions has been largely pragmatic and based on typical methods used in modern articulatory models. Further analysis of the sensitivity of the transformation to landmark choice and smoothing process should be undertaken. The VT area functions are calculated using the male adult alpha-beta from

Soquet *et al.*'s [57] paper. In the absence of soft tissue data from Neanderthals, this seems a reasonable choice, although other parameter sets could be used if preferred. As well as the fixed parameters, there are a number of user-selectable parameters: in particular, relating to how bony landmarks are grouped and how sections of the MH VT are associated with them to define their transformation vectors. There is scope for a systematic study of the effect of different groupings and affiliations on the estimates of the formants.

The options for further exploration with the Simus_Neanderthals model are wide. We envisage studies with meshes from scans of other Neanderthal fossils, and systematic studies of the effects of different combinations of modelling assumptions and parameter choices on VT acoustic outputs. We should emphasize that we view the model as a tool for hypothesis testing rather than as a definitive answer to the question of Neanderthal VT anatomy. It offers the opportunity for an incremental approach to the question of speech potential in fossil hominins and moves away from the more categorical approach found in much of the historical literature.

in a grant to the HANDTOMOUTH project (FP6, Contract no. 29065, NEST-2004-PATH-HUMAN).

ENDNOTE

[1]The fossil specimens matched with the non-human primate/neonatal human/Neanderthal condition included australopithecines (*Australopithecus africanus*, *Paranthropus robustus*, *Paranthropus boisei*), and also other classic Neanderthal specimens (Saccopastore 1, Monte Circeo, Teshik-Tash infant, La Ferrassie 1, La Quina infant, Pech-de-l'Azé, Shanidar 1), as well as Solo 2. The fossil specimens matched with the adult human condition included Steinheim, Broken Hill, Skhul 5, Djebel Qafzeh and Cro-Magnon ([29], p. 83).

REFERENCES

1 Arbib, M. A. 2006 Aphasia, apraxia and the evolution of the language-ready brain. *Aphasiology* **20**, 1125–1155. (doi:10.1080/02687030600741683)

2 Roy, A. & Arbib, M. 2005 The syntactic motor system. *Gesture* **5**, 7–37. (doi:10.1075/gest.5.1.03roy)

3 Kirby, S. 2000 Syntax without natural selection: how compositionality emerges from vocabulary in a population of learners. In *The evolutionary emergence of language: social function and the origins of linguistic form* (ed. C. Knight), pp. 303–323. Cambridge, UK: Cambridge University Press.

4 Kirby, S., Dowman, M. & Griffiths, T. 2007 Innateness and culture in the evolution of language. *Proc. Natl Acad. Sci. USA* **104**, 5241–5245. (doi:10.1073/pnas.0608222104)

5 Corballis, M. C. 2003 Out on a limb. *Laterality* **8**, 195–200.

6 Corballis, M. C. 2003 From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sciences* **26**, 199–260.

7 Steele, J. & Uomini, N. T. 2005 Humans, tools and handedness. In *Stone knapping: the necessary conditions for a uniquely hominin behaviour* (eds V. Roux & B. Bril), pp. 217–239. Cambridge, UK: McDonald Institute for Archaeological Research.

8 Steele, J. & Uomini, N. 2009 Can the archaeology of manual specialization tell us anything about language evolution? A survey of the state of play. *Camb. Archaeol. J.* **19**, 97–110. (doi:10.1017/S0959774309000067)

9 Uomini, N. T. 2010 Handedness in Neanderthals. In *Neanderthal lifeways, subsistence and technology* (eds N. J. Conard & J. Richter). Heidelberg, Germany: Springer.

10 Arensburg, B., Tillier, A. M., Vandermeersch, B., Duday, H. & Rak, Y. 1989 A middle Palaeolithic human hyoid bone. *Nature* **338**, 758–760. (doi:10.1038/338758a0)

11 Martínez, I., Arsuaga, J.-L., Quam, R., Carretero, J.-M., Gracia, A. & Rodríguez, L. 2008 Human hyoid bones from the middle Pleistocene site of the Sima de los Huesos (Sierra de Atapuerca, Spain). *J. Hum. Evol.* **54**, 118–124. (doi:10.1016/j.jhevol.2007.07.006)

12 Maclarnon, A. & Hewitt, G. 2004 Increased breathing control: another factor in the evolution of human language. *Evol. Anthropol.* **13**, 181–197. (doi:10.1002/evan.20032)

13 Krause, J. *et al.* 2007 The derived *FOXP2* variant of modern humans was shared with Neanderthals. *Curr. Biol.* **17**, 1908–1912. (doi:10.1016/j.cub.2007.10.008)

14 Gibson, K. R. & Ingold, T. (eds) 1993 *Tools, language and cognition in human evolution*. Cambridge, UK: Cambridge University Press.

15 Greenfield, P. M. 1991 Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**, 531–550. (doi:10.1017/S0140525X00071235)

16 Ambrose, S. H. 2001 Paleolithic technology and human evolution. *Science* **291**, 1748–1753. (doi:10.1126/science.1059487)

17 Camps, M. & Uriagereka, J. 2006 The Gordian Knot of linguistic fossils. In *The biolinguistic turn. Issues on language and biology* (eds J. Rosselló & J. Martín), pp. 34–65. Barcelona, Spain: Universitat de Barcelona.

18 Pastra, K. & Aloimonos, Y. 2011 The minimalist grammar of action. *Phil. Trans. R. Soc. B* **367**, 103–117. (doi:10.101098/rstb.2011.0123)

19 Stout, D. & Chaminade, T. 2011 Stone tools, language and the brain in human evolution. *Phil. Trans. R. Soc. B* **367**, 75–87. (doi:10.101098/rstb.2011.0099)

20 Lieberman, P. 1975 *On the origins of language*. New York, NY: Macmillan.

21 Fant, G. 1960 *Acoustic theory of speech production*, 2nd edn. The Hague, The Netherlands: Mouton.

22 Stevens, K. N. 1989 On the quantal nature of speech. *J. Phonetics* **17**, 3–46.

23 Apostol, L., Perrier, P. & Bailly, G. 2004 A model of acoustic interspeaker variability based on the concept of formant-cavity affiliation. *J. Acoust. Soc. Am.* **115**, 337–351. (doi:10.1121/1.1631946)

24 Lieberman, P. & Crelin, E. S. 1971 On the speech of the Neanderthal man. *Linguist. Inquiry* **2**, 203–222.

25 Lieberman, P. H., Crelin, E. S. & Klatt, D. H. 1972 Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee. *Am. Anthropol.* **74**, 287–307. (doi:10.1525/aa.1972.74.3.02a00020)

26 Henke, W. L. 1966 *Dynamic articulatory model of speech production using computer simulation*. PhD thesis, MIT, Cambridge, MA.

27 Peterson, G. E. & Barney, H. L. 1952 Control methods used in a study of vowels. *J. Acoust. Soc. Am.* **24**, 175–184. (doi:10.1121/1.1906875)

28 Lieberman, P. & Crelin, E. S. 1971 On the speech of Neanderthal man. *Linguistic Inquiry* **2**, 203–222.

29 Lieberman, P. 1973 On the evolution of language: a unified view. *Cognition* **2**, 59–94.

30 Takemoto, H. 2001 Morphological analyses of the human tongue musculature for three-dimensional modeling. *J. Speech Lang. Hear Res.* **44**, 97–107. (doi:10.1044/1092-4388(2001/009))

31 Takemoto, H. 2008 Morphological analyses and 3D modeling of the tongue musculature of the chimpanzee (*Pan troglodytes*). *Am. J. Primatol.* **70**, 966–975. (doi:10.1002/ajp.20589)

32 Nishimura, T., Mikami, A., Suzuki, J. & Matsuzawa, T. 2006 Descent of the hyoid in chimpanzees: evolution of face flattening and speech. *J. Hum. Evol.* **51**, 244–254. (doi:10.1016/j.jhevol.2006.03.005)

33 Nishimura, T., Mikami, A., Suzuki, J. & Matsuzawa, T. 2007 Development of the laryngeal air sac in chimpanzees. *Int. J. Primatol.* **28**, 483–492. (doi:10.1007/s10764-007-9127-7)

34 Nishimura, T., Oishi, T., Suzuki, J., Matsuda, K. & Toshimitsu, T. 2008 Development of the supralaryngeal vocal tract in Japanese macaques: implications for the evolution of the descent of the larynx. *Am. J. Phys. Anthropol.* **135**, 182–194. (doi:10.1002/ajpa.20719)

35 Nishimura, T. 2003 Comparative morphology of the hyo-laryngeal complex in anthropoids: two steps in the evolution of the descent of the larynx. *Primates* **44**, 41–49. (doi:10.1007/s10329-002-0005-9)

36 Nishimura, T. 2005 Developmental changes in the shape of the supralaryngeal vocal tract in chimpanzees. *Am. J. Phys. Anthropol.* **126**, 193–204. (doi:10.1002/ajpa.20112)

37 Laitman, J. T., Heimbuch, R. C. & Crelin, E. S. 1979 The basicranium of fossil hominids as an indicator of

their upper respiratory systems. *Am. J. Phys. Anthropol.* **51**, 15–34. (doi:10.1002/ajpa.1330510103)

38 Laitman, J. & Riedenberg, J. 1993 Specializations of the human upper respiratory and upper digestive systems as seen through comparative and developmental anatomy. *Dysphagia* **8**, 318–325. (doi:10.1007/BF01321770)

39 Lieberman, D. E. & McCarthy, R. C. 1999 The ontogeny of cranial base angulation in humans and chimpanzees and its implications for reconstructing pharyngeal dimensions. *J. Hum. Evol.* **36**, 487–517. (doi:10.1006/jhev.1998.0287)

40 Lieberman, D. E., McCarthy, R. C., Hieemae, K. M. & Palmer, J. B. 2001 Ontogeny of postnatal hyoid and larynx descent in humans. *Arch. Oral Biol.* **46**, 117–128. (doi:10.1016/S0003-9969(00)00108-4)

41 Boë, L.-J., Heim, J.-L., Honda, K. & Maeda, S. 2002 The potential Neanderthal vowel space was as large as that of modern humans. *J. Phonetics* **30**, 465–484. (doi:10.1006/jpho.2002.0170)

42 Maeda, S. 1989 Compensatory articulation during speech: evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and modelling* (eds W. J. Hardcastle & A. Marchal), pp. 131–149. Dordrecht, The Netherlands: Kluwer Academic.

43 Boë, L., Heim, J., Honda, K., Maeda, S., Badin, P. & Abry, C. 2007 The vocal tract of newborn humans and Neanderthals: acoustic capabilities and consequences for the debate on the origin of language. A reply to Liberman (2007a). *J. Phonetics* **35**, 564–581. (doi:10.1016/j.wocn.2007.06.006)

44 Badin, P. & Fant, G. 1984 Notes on vocal tract computations. *KTH Speech Transm. Lab. Q. Prog. Status Rep.* **2–3**, 53–108.

45 Heim, J.-L. 1990 La nouvelle reconstitution du crâne néandertalien de la Chapelle-aux-Saints. Méthode et résultats. *Bull. Mém. Soc. Anthropol. Paris* **6**, 94–117.

46 de Boer, B. & Fitch, W. T. 2010 Computer models of vocal tract evolution: an overview and critique. *Adapt. Behav.* **18**, 36–47. (doi:10.1177/1059712309350972)

47 Maeda, S. 1989 Compensatory articulation during speech: evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and modelling* (eds W. J. Hardcastle & A. Marchal), pp. 131–149. Boston, MA: Kluwer Academic.

48 de Boer, B. 2009 Why women speak better than men (and its significance for evolution). In *The prehistory of language* (eds R. Botha & C. Knight), pp. 255–265. Oxford, UK: Oxford University Press.

49 de Boer, B. 2010 Investigating the acoustic effect of the descended larynx with articulatory models. *J. Phonetics* **38**, 679–686. (doi:10.1016/j.wocn.2010.10.003)

50 Mermelstein, P. 1973 Articulatory model for the study of speech production. *J. Acoust. Soc. Am.* **53**, 1070–1082. (doi:10.1121/1.1913427)

51 Kelly Jr, J. L. & Lochbaum, C. C. 1973 Speech synthesis. In *Speech synthesis* (eds J. L. Flanagan & L. R. Rabiner), pp. 127–130. Stroudsburg, PA: Dowden, Hutchison & Ross, Inc.

52 Tilotta, F., Richard, F., Glaunès, J., Berar, M., Gey, S., Verdeille, S., Rozenholc, Y. & Gaudy, J.-F. 2009 Construction and analysis of a head CT-scan database for craniofacial reconstruction. *Forensic Sci. Int.* **191**, 112.e111–112.e112. (doi:10.1016/j.forsciint.2009.06.017)

53 Rohlf, F. J. 2006 tpsDig *v. 2.10*. See http://life.bio.sunysb.edu/morph/soft-dataacq.html.

54 Stewart, T. D. 1962 Neanderthal cervical vertebrae—with special attention to the Shanidar Neanderthals from Iraq. *Bibl. Primatol.* **1**, 130–154.

55 Gómez-Olivencia, A., Carretero, J. M., Arsuaga, J. L., Rodríguez-García, L., García-González, R. & Martínez, I. 2007 Metric and morphological study of the upper cervical spine from the Sima de los Huesos site (Sierra de Atapuerca, Burgos, Spain). *J. Hum. Evol.* **53**, 6–25. (doi:10.1016/j.jhevol.2006.12.006)

56 Badin, P. & Serrurier, A. 2006 Three-dimensional linear modeling of tongue: articulatory data and models. In *Proc. 7th Int. Seminar on Speech Production, Ubatuba, Brazil*, 13–15 December 2006, pp. 395–402.

57 Soquet, A., Lecuit, V., Metens, T. & Demolin, D. 2002 Mid-sagittal cut to area function transformations: direct measurements of mid-sagittal distance and area with MRI. *Speech Commun.* **36**, 169–180. (doi:10.1016/S0167-6393(00)00084-4)

58 Maeda, S. 1982 A digital simulation method of the vocal-tract system. *Speech Commun.* **1**, 199–229. (doi:10.1016/0167-6393(82)90017-6)

59 Liljencrants, J. & Lindblom, B. 1972 Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* **48**, 839–862.

60 Lieberman, D. E. 2008 Speculations about the selective basis for modern human craniofacial form. *Evol. Anthropol.* **17**, 55–68. (doi:10.1002/evan.20154)

61 Trinkaus, E. 2003 Neanderthal faces were not long; modern human faces are short. *Proc. Natl Acad. Sci. USA* **100**, 8142–8145. (doi:10.1073/pnas.1433023100)

62 Badin, P., Perrier, P., Boë, L.-J. & Abry, C. 1990 Vocalic nomograms: acoustic and articulatory considerations upon formant convergences. *J. Acoust. Soc. Am.* **87**, 1290–1300. (doi:10.1121/1.398804)

63 Sundberg, J. 1977 Acoustics of the singing voice. *Sci. Am.* **236**, 82–100. (doi:10.1038/scientificamerican0377-82)

Research

The minimalist grammar of action

Katerina Pastra[1,*] and Yiannis Aloimonos[2]

[1]*Cognitive Systems Research Institute, 7 Makedonomachou Prantouna Street, Athens 11525, Greece*
[2]*Computer Vision Laboratory, University of Maryland, College Park, MD 20742, USA*

Language and action have been found to share a common neural basis and in particular a common 'syntax', an analogous hierarchical and compositional organization. While language structure analysis has led to the formulation of different grammatical formalisms and associated discriminative or generative computational models, the structure of action is still elusive and so are the related computational models. However, structuring action has important implications on action learning and generalization, in both human cognition research and computation. In this study, we present a biologically inspired generative grammar of action, which employs the structure-building operations and principles of Chomsky's Minimalist Programme as a reference model. In this grammar, action terminals combine hierarchically into temporal sequences of actions of increasing complexity; the actions are bound with the involved tools and affected objects and are governed by certain goals. We show, how the *tool role* and *the affected-object role* of an entity within an action drives the derivation of the action syntax in this grammar and controls recursion, merge and move, the latter being mechanisms that manifest themselves not only in human language, but in human action too.

Keywords: generative grammar of action; tool use; action syntax; action decomposition; temporal sequence; minimalist grammar

1. INTRODUCTION

The repertoire of human actions is infinite, starting from the simplest intentional body movements such as *stretching a leg* to creative *dancing routines*, to interaction with tools and objects such as *grasping a knife*, to even more complex series of actions that formulate events, such as *preparing a salad* or *cleaning the house*. Uncovering the structure of action has been a quest in many disciplines, including cognitive science and artificial intelligence. How could one generate or parse actions of any complexity avoiding at the same time overgeneralization? The question is similar to an analogous problem in language analysis: How could one generate or parse all and only the grammatical sentences of a language?

The quest for the structural principles of visual and motoric action goes back at least to the early fifties and suggestions made by the psychologist Lashley [1] that syntax may apply not only to language but also to other forms of behaviour, such as goal-directed action. From another perspective, the archaeologist Andrè Leroi-Gourhan argued that bipedality led to technology and technology (tool making and use) reflects a capability (for derivation of structures) that may link human action and language [2]. Since then, corroborating experimental evidence on the relation between action and language and the hierarchical structure of

action, in particular, abounds; for example, two-year-old children have been found to be able not only to parse hierarchically organized actions [3], but also to copy and reproduce such actions [4]. Complex action structure (analysed as means-end parse trees) has also been found to be represented abstractly, i.e. independently of the actual semantics of the actions [5]. More strikingly, neurobiological evidence on the nature of neural circuits in the traditionally related to language-production area of the human brain (i.e. Broca's area) provides a growing number of suggestions regarding the characteristics of an action grammar, such as the role of body parts/effectors, of tools and object type, and the role of the notion of 'goal' in human action representation [6,7].

However, specifying an action grammar that will generate thousands of actions is still elusive. There are only very few attempts for developing an action grammar in computational research [8–10] and these are recognitive rather than generative approaches. There is a need for developing a generative grammar of action that will have both computational expressivity and simplicity, and a biological basis; the former will allow for employing the grammar in artificial intelligence applications, while the latter may prove to be the key for action learning and generalization.

In this study, we employ a formal language analysis framework as a reference model for presenting a generative grammar of action. In particular, we employ the generative grammar framework, for crossing over human language, to human action. Although there is a variety of grammars for describing the structure of language, we choose the Chomskyan approach and

* Author for correspondence (kpastra@csri.gr).

its latest evolution into the minimalist programme (MP) [11], because it is the culmination of an attempt to describe and explain language syntax in terms of more general principles and operations that are not tightly tied to the idiosyncrasies of the human language system, but instead may have counterparts in other biological systems [12]. This perspective allows one to look for universals not only within the structures of different human languages, but also across natural language to non-symbolic sensorimotor spaces, such as human action.

We present the characteristics and components of this grammar of action, many aspects of which are corroborated by neurobiological findings. We argue that the notion of 'tool use' drives action syntax derivation and through examples, we present how this takes place when employing the suggested action grammar.

2. RESEARCH ON THE STRUCTURE OF ACTION

The structure of visual and motoric action has been explored by a number of disciplines, including neuroscience, psychology, computer vision and robotics. In this section, we make a concise presentation of neurobiological and computational research related to the existence and implementation of a grammar of action.

(a) *Neurobiological approaches on action grammar*

Recent years have seen an increasing body of experimental evidence suggesting that Broca's area, the human brain area traditionally linked to language production, is involved in representing complex hierarchical structures regardless of modality, such as those involved in action execution and observation [13]. In other words, Broca's area has been suggested as the neural locus of an action grammar [14], an area where goals are represented and hierarchical motor chains are planned [7]. The findings indicate a common syntactic-like structure between language and action that has led to speculations that 'this capacity evolved from motor and premotor functions associated with action execution and understanding such as those characterizing the mirror neurons' [13].

At a behaviour level, action syntax has been shown to comprise simpler elements (motor primitives) that are connected to each other either serially or in parallel (i.e. simultaneously; [15–19]). Researchers have concentrated on the analysis of many different actions, such as reaching and grasping, gait and balance, posture and locomotion. Reaching movements appear to be coded in terms of direction and extent, and appear to be composed of discrete submovements, all with a similar stereotypical, serially concatenated shape and overlapping in time [20–22]. Human and monkey grasping and object manipulation has been studied extensively and has been described as consisting of sub-actions executed as a unified coordinated complex act (e.g. [23]). Parallel syntax, on the other hand, involves the simultaneous activation of several muscles that produce a torque about a joint or a force in a particular direction. Electromyogram recordings from frog hind limb muscles have been analysed to test whether natural behaviour shows synergies among groups of muscle activities for an entire set of natural behaviours [18,24–26]. Similar attempts have been made to find muscle synergies during human posture and locomotion [27,28].

In some approaches, motor primitives basically amount to motor schemas or control modules that may be specific to a task; for example, in the 'motor ideas/schemas' approach, coordinated control programmes regulate coactivation of perceptual and motor schemas and the passing of action parameters from one to another to determine hand–environment interaction [29,30]. Within this approach, perceptual schemas represent objects that are involved in an action, while motor schemas represent the actual motor programme to be executed.

Sequential or parallel, the combination of action primitives or schemas into complex actions has been explored, but has not led yet to a grammar that will allow one to generate thousands of actions, incorporating the ever-growing body of related biological evidence.

(b) *Computational grammars of action*

At a computational level, there is really not much previous work on the subject, i.e. on a computational *motoric grammar* for action. A system that comes closest in spirit to a grammar for action was developed in Juhola [31] more than 15 years ago for handling eye movements. By turning the eye movement data into a string of symbols, they developed a finite automaton (the equivalent of a regular grammar) for representing the data. However, some researchers have come close to the idea of motoric primitives of action and primitives are, indeed, the first step to a grammar. A number of data transformations have been employed to derive a limited number of motor primitives that are then combined through a well-defined set of rules to form more complex actions (see, for example, the movements of del Vecchio *et al.* [32] or the modules of Jenkins & Mataric [33]). Primitives in these cases may be kinematic, dynamic or kinematodynamic [16,18,34,35], and are extracted using statistical techniques such as principal component analysis or hidden Markov models (HMM), and others.

In a recognitive (rather than generative) approach to action analysis, decomposition of action sequences into primitives has taken many forms. Finger movements and forces have been decomposed into basic synergies based either on the idea of uncontrolled manifold or on inverse dynamics computations [35,36]. Hand gestures also consist of primitives or more complicated sequences that can be decomposed into a series of more elementary units of activity [37]. In Fod *et al.* [38], primitives were extracted by *k*-means clustering the projection of high-dimensional segment vectors onto a reduced subspace, while in Kahol *et al.* [39] the local minimum in total body force was used to detect segment boundaries. In Nakazawa *et al.* [40], similarities of motion segments were measured according to a dynamic programming distance and clustered with a nearest-neighbour algorithm. In Wang *et al.* [41], gestures were segmented with the local minima of velocity and local maxima of change in direction. These segments were

hierarchically clustered into classes using HMM to compute a metric. Grammar induction techniques were applied to both motion capture data and images (silhouettes) to produce a human activity language [10,42], thus formalizing and unifying several prior approaches.

The development of a generative grammar for action, i.e. one that can be used both for visual action analysis and for generation of goal-directed behaviour, is of primary importance for both computer vision and robotic applications. Clearly, such grammar is not only missing in state-of-the-art computational approaches to action analysis, but is also elusive at a formal analysis (theoretical) level. This is the contribution of our study: the development of a generative grammar of action and in particular, of a grammar with computational applicability and with biological bases, the latter being for us a prerequisite for scalability and generalization of a computational approach.

3. HOW IS ACTION STRUCTURED?

In order to answer this question, we employ a formal analysis framework that has been developed for language. It is the latest formulation of the Chomskyan tradition of generative grammars, the MP [11]. The MP and the generative grammar paradigm in general have, indeed, many details and intricacies for dealing with a number of phenomena in language. There are many ways of implementing the theory and representing information in the parse trees, with versions of the theory before the MP being implemented and elaborated more extensively. In this study, we do not wish to go into the details of the representation and the theory, or to follow strictly one or another approach in parse tree representation. Our aim is to present the basic framework, so that it becomes obvious how we employ it as a reference model for developing a grammar of action. Therefore, in this section, we will first introduce this formal analysis approach and then we will present our use of the principles and syntactic operations described in the framework to formulate a generative grammar of action.

(a) *The Chomskyan tradition of generative grammars*

Generative grammars have been used extensively for the analysis of the structure of human language. Simply put, a generative grammar comprises a set of elements and a set of production (rewrite) rules that correctly predict which combinations of elements form grammatical sentences. A particular type of generative grammars are the phrase structure grammars or else *context-free grammars*, which have recursive rules, i.e. they allow for nesting of elements in same type elements, accommodating thus for embedded structures. These grammars comprise a set of terminals (e.g. lexical categories such as noun, verb, adjective), a set of non-terminals (i.e. the phrases, such as noun phrase, verb phrase, etc.) and a set of production rules of the form $X \rightarrow y$, where X is a single non-terminal symbol, and y is a string of zero or more terminals and/or non-terminals. The context of X within a structure does not affect the use of the

corresponding rule (hence context-free). In applying the grammar for the analysis of a certain structure, a parse tree is produced, in which non-terminal symbols are the nodes, terminal symbols are the leaves and each node expands (through successive application of the production rules) into the next level of the tree [43].

Although highly expressive, this type of grammar cannot account for natural language phenomena such as agreement (e.g. case, number, gender agreement) and reference (e.g. anaphora, relative clauses). These are cases of either 'discontinuous elements' or long-distance dependencies between constituents of a sentence [44,45]. The Chomskyan tradition of *generative grammar* deals with such phenomena through the use of a number of processes (transformations) on the output of context-free grammars [46]. The latest evolution of the Chomskyan grammar tradition is the MP [11], a framework that reduces transformation grammar to a simple, powerful computational mechanism imbued with the principle of economy/minimalism in both derivation and representation of syntactic structures; this minimalism advocates that: (i) *minimal derivation processes run* for producing the syntactic structure (only those transformations needed to fully interpret the constituents of the structure) and (ii) *minimal representations of syntactic structures are produced* (only what is needed to satisfy grammaticality).

The language that has a generative grammar consists of:

— a finite set of terminals T, i.e. leaf nodes in a parse tree, or else minimal projections, the actual lexical units that make up a sentence; in the MP, these are characterized through a number of morpho-syntactic features F, such as their part of speech, case, type of complement, etc.;
— a finite set of non-terminals NT, i.e. phrase types, syntactic categories of the terminals such as noun phrases, verb phrases, etc. such that T ∩ NT = ø; and
— a finite set of production rules R, i.e. rewrite rules that are applied to terminals and non-terminals recursively (i.e. a rule rewrites as a previous rule or as itself) producing an infinite number of grammatical structures (see table 1).

The generative grammar rules in table 1 work as follows. X is the minimal projection of a language unit (the actual word), none of the features of which have been checked (has been attributed a value). X′ is an intermediate level projection in which some of the features of the unit have been checked, and X″ is the maximal projection of the unit, in which all its features have been checked. Y is a specifier when it precedes X, i.e. a terminal or non-terminal that modifies the meaning of X, and a complement when it follows X, i.e. a terminal or non-terminal that complements the meaning of X. Parenthesis denotes that its presence is optional. Table 1 shows that such grammars can be used for the derivation of complex language structures involving sentences within sentences (second column), which may be

Table 1. Generative grammar rules.

general case	instantiation for sentences	instantiation for noun phrases	instantiation for verb phrases
$X'' \rightarrow (y), X'$	$S'' \rightarrow (y), S'$	$N'' \rightarrow (y), N'$	$V'' \rightarrow (y), V'$
$X' \rightarrow y, X'$	$S' \rightarrow y, S'$	$N' \rightarrow y, N'$	$V' \rightarrow y, V'$
$X' \rightarrow X', y$	$S' \rightarrow S', y$	$N' \rightarrow N', y$	$V' \rightarrow V', y$
$X' \rightarrow X, (y)$	$S' \rightarrow S, (y)$	$N' \rightarrow N, (y)$	$V' \rightarrow V, (y)$

analysed down to the level of noun and verb phrases forming a sentence. The structure of noun and verb phrases themselves is also analysed in terms of maximal projections (third and fourth column, respectively).

In the MP framework, the derivation of a syntactic structure starts *bottom-up*; a *Merge* function checks the features of a terminal (lexical unit) and for those features with un-attributed values (i.e. variables), it initiates a *Search* for another unit whose feature-values can be unified with the variables. This merging creates binary structures and is applied recursively until all features are 'interpreted' (have a value) [47]. For example, in a simple request such as 'grasp the knife', the verb 'grasp' has an object-complement feature object complement (O_c) with category type 'nominal', case 'accusative' and semantic type 'graspable object'; merging initiates a search for a lexical item with these features, i.e. [+nominal, +accusative, +graspable] in order to fill in the object-complement variable O_c. The determiner 'the' initiates a further search for satisfying its own features, which leads to the creation of the noun phrase 'the knife'. This noun phrase can now satisfy the 'grasp' search for an element or structure that interprets its own feature variables; so a further merging takes place. Every merged set of elements (phrase) has a label (the head of the phrase) that determines the properties of the phrase, e.g. {the {the, knife}}, {grasp {grasp, the knife}}. These properties allow certain projections and eliminate alternatives.

One form of merging is the *Move* operator. It is 'merging' of one morpho-syntactic element with itself (internal merging). For internal merging to take place, a probe–goal relation must hold between at least one feature of an element and a corresponding feature of another element [11]. For example, in the sentence 'which knife$_i$ has John grasped $\emptyset_i$' there is an internal merging between the moved element 'knife' (normally expected after the verb, where the 'null' element with trace 'i' is) and its co-indexed trace (i).

(b) *A minimalist grammar of action*

In employing a generative grammar for the analysis of the structure of human action, one needs to define the set of terminals, features, non-terminals and production rules in the sensorimotor domain. So, which are these terminals and non-terminals, which are their 'morpho-syntactic' features and how do they merge creating more and more complex actions?

In what follows, we will present a minimalist grammar of action that consists of action primitives (terminals), action 'phrases' (non-terminals) and

three action features that drive merging, namely the *tool complement* of an action, the *affected-object* complement and the *goal* of the whole action structure.

In our analysis, we consider a human action to be a *serial or parallel conjunction of perceptible movements carried out by one or more actors with a certain goal*. We identify three main 'morpho-syntactic' features that characterize human actions and that we employ for defining action terminals and non-terminals. These 'morpho-syntactic features' are actually parameters that affect the execution of actions and distinguish one *action type* from another; they go beyond movement execution features (e.g. direction, velocity, etc.) that they modify though, as soon as a movement is embedded within action context. We will go through these parameters, referring in parallel to an example action parse tree shown in figure 1.

(i) *Tool complement* (t_c)

This is the effector of a movement, this being a body part, a combination of body parts or the extension of a body part with a graspable object used as a tool.

Actions are always being executed through the use of an effector (body part) or its extension, i.e. an instrument (artefact). Being a body part or artefact, the 'tool' used changes the execution of the action in terms of configuration of the effector, force exerted, etc. For example, grasping something with the hand is different (in its motoric execution and complexity) from grasping it with pliers, which is different from grasping it with tweezers and so on. What we refer to here is a merging in the motoric space, in which any action necessarily/inherently requires a 'tool complement' so, a search for the entity that interprets this variable is initiated for deriving a first binary action structure. See for example, the action-tool binary branches of the tree in figure 1.

Any graspable entity can be used as a tool in the realization of an action (e.g. use of a book to pound something). In some cases, the particular use of the entity is common (e.g. use of a hammer for pounding); in other cases it may be uncommon, but still possible (e.g. the case of book for pounding). We consider this an essential feature of any human action, which is syntactic, i.e. it is explicitly present in the perception or execution of an action, as an independent constituent.

We have to note that under the notion of a 'tool complement', we include both body parts and artefacts/instruments, suggesting essentially that our body part effectors are tools (means for performing actions) in the same way as other objects/artefacts may be used as 'means' for achieving a task. This is corroborated by neurobiological experiments that indicate that tools are indeed perceived as extensions of one's own body part [48,49]; so there is an intricate relation between body parts and other objects through the attribution of a 'tool role' to them. Recent experimental findings have also shown that in visual information processing, humans differentiate graspable objects consistently faster than non-graspable ones, and among graspable ones, prototypical tools (e.g. hammer) are differentiated faster than natural

kinds (e.g. carrot) [50]; more importantly, this differentiation was found to be mediated by the activation of motor areas (cf. also [6], on the visuomotor neurons). This evidence suggests that the 'toolness' of an object may be an important differentiation parameter in an action grammar.

(ii) *Object complement* (O_c)

This is any object affected by a tool use action. This is another syntactic feature of action; it is the object of interaction, to which the effects of an action are transferred. It may be any entity. This feature actually calls for a further merge operation, between the action-tool structure and the affected object. It results in {action-tool, object} binary structures such as the ones shown in figure 1 between branches of the tree related through the action-object relation, e.g. {grasp with hand$_1$, knife}. The object that is affected by an action differentiates the action itself; for example, *grasping a pencil with the hand* is different from *grasping a glass with the hand*, not only because the goal may be different but also because of the characteristics of the object grasped. Biological evidence of strict congruence of action type with object type (e.g. discharge of 'precision grip neurons' when small objects are observed in canonical visuomotor neurons in area F5) [6] suggests that object complements are indeed differentiation parameters for actions.

(iii) *Goal* (g)

This is the final purpose of an action sequence of any length or complexity. Another important feature of actions that modify their execution is their goal. This is a morphological feature, i.e. it is not an explicit, independent syntactic constituent of the action phrase, but instead an 'inflectional' feature, a parameter that modifies the execution of the constituents of an action in terms of effector configuration and spatial interaction with tool and object complements. The execution of a movement is modified according to the action sequence in which it is embedded, i.e. according to the final goal of the action. An analogy to the phenomenon of agreement in sentences can be drawn here: person and number agreement of words, for example, modify the words of a sentence themselves; for a grammatical sentence, all words must be in agreement. Similarly, in action, all sub-actions must agree in terms of the final goal to be served. Thinking of a word or phrase stripped of any agreement indicators is as artificial as thinking of a purposeless action. This is related to what Luria called 'kinetic melody' [51], i.e. the fluidity of motor acts as they follow one another.

Going back to the example in figure 1, in the action parse tree, the goal feature is attached to a null-constituent and dominates the whole action sequence and its sub-actions. In producing the action tree bottom up (i.e. as the action evolves in time), the goal feature of the sub-actions remains unspecified; it is only when all other features have been checked and no more actions can be merged into a common complex structure that the goal feature can be checked

too, i.e. when the top node of the action tree is reached; the goal is the final action.

This role of the goal feature in our grammar is supported by experimental findings that show that certain neurons discharge only during goal-directed action [6], and many of them have categorization, generalization and specificity characteristics—for example, they go beyond effector differences [6,52]. In Fogassi *et al.* [7], it was shown in the monkey brain that certain neurons go beyond object type differences when the same movements share a goal, and that the intention of an action sequence is reflected already in its first sub-action [7]. The latter was also shown to be the case in the human brain; in Cattaneo *et al.* [53], a forward activation of motor sequences was shown in typically developing children (as opposed to children with autism). In these experiments, increased activity of the muscles involved in mouth opening was found before grasping takes place in 'grasping to eat' action sequences (versus grasping to displace ones), during both action observation and action execution; such activation denotes that the final goal of the action sequence was predicted (and actually 'experienced') from the very first phases of the sequence.

In other words, these findings point to important aspects of the role played by the final goal of an action structure:

— the realization of the same movement type with the same tool and object complements changes when the goal of the action changes, e.g. *grasping a pencil in order to displace it* is different from *grasping a pencil in order to write*;

— the realization of the same movement type with different tool and/or affected object changes, even if the goal of the action remains the same, e.g. *grasping an apple to displace it* is different from *grasping a cube to displace it* (though in such cases effects of the expected/common goal of an object seem to be present, cf. [7]); and last

— the final goal of an action sequence is predicted from the very first sub-action(s) of the sequence; for example, sub-components of the grasping activity such as to '*extend the hand towards the pencil*' involve a configuration of the effector that depends on the final goal. Actually, the corresponding 'grasping neurons' begin to discharge before the object-hand contact [6], while the pre-shaping of the fingers also takes place during the transfer of the hand [30].

We need to note that this 'goal' feature that governs all constituents of an action structure is the *global goal* (the final goal) of the action structure. One may argue that each sub-action of an action structure may have its own *local/immediate goal* too; for example, in figure 1, extending hand$_1$ towards something has the immediate goal of enclosing in hand$_1$ this object (i.e. grasping it); its immediate goal is the next action that it enables. In artificial intelligence, traditional planning techniques for the analysis/execution of a task divide the task into sub-goals, i.e. into steps with their own immediate goals. However, the neurobiological evidence mentioned earlier points to the fact that the final goal of
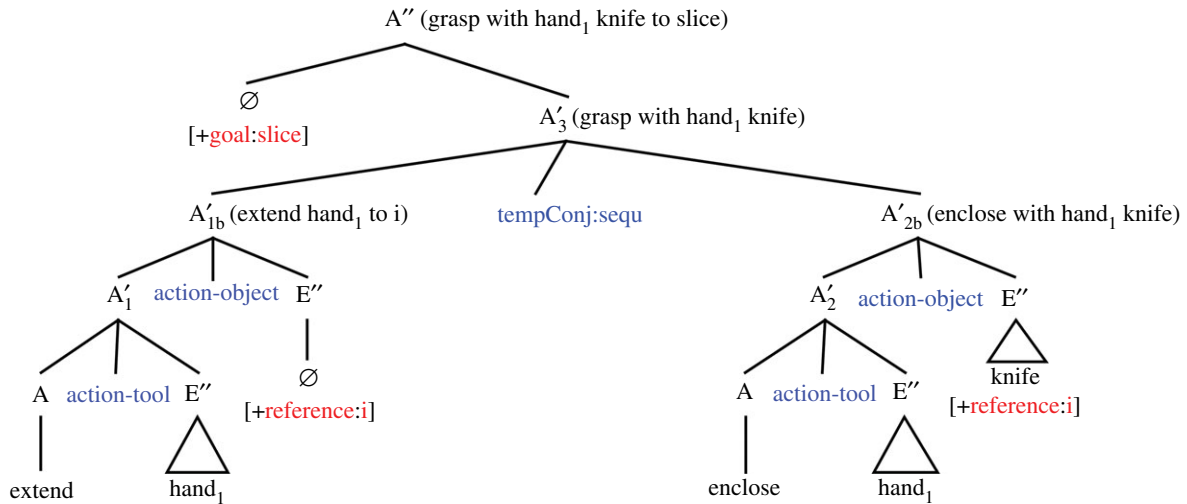
Figure 1. Part of an action tree for 'grasp with hand$_1$ knife to slice'; A stands for action primitives (terminals), A′ for action structures (non-terminals), A″ for the maximal projection of an action structure. E″ is the maximal projection of an entity structure. Triangles in the tree denote that the corresponding part of the tree is not fully analysed for keeping the figure simple. Parentheses present the morphological features of the corresponding tree nodes, in an 'attribute:value' format; the plus sign denotes the presence of such features, and a minus would denote the absence of a feature. The exact type of relation between branches of the tree is clearly denoted for clarification purposes; 'action-tool' and 'action-object' are complements of an action and as such they are inherently related to the corresponding action structure. Sub-actions of a complex action are sequential or parallel in time, i.e. they are related through the corresponding 'temporal conjunction' type (tempConj:sequ, or tempConj:par).

a complex action is evident even in its very first phases; it affects the motoric execution of the sub-actions and it is evident in the early activation of muscles that are related to final action constituents of the action sequence. As shown in the next section, our minimalist grammar of action makes no use of sub-goals; instead, it is the final goal of an action structure that is required for deriving the maximal projection of an action (i.e. A″ in the action tree).

Other 'morpho-syntactic' features of action are *modifiers* that denote the location/scene an action takes place at, or an object that is used as the location of an action (e.g. 'slice bread *on the table*'); these do not inherently affect the execution of the action itself, and they specify the setup of the action. So, their presence is optional.

The tool and object complements as well as the modifiers are entities; these entities have their own perceptual (e.g. visual or other) grammar, the terminals and non-terminals of which could be defined in terms of the action grammar. Elaboration on a perceptual (e.g. visual) grammar of objects is beyond the scope of this paper; however, some general definitions should be in place:

Entity terminals
These are the simplest entities (objects) that can be defined as perceptible entities that participate in at least one motor programme and do not comprise other entities themselves. They are distinguished from each other through their perceptible features (colour, shape, texture, etc.) and the role they play in the motor programmes in which they participate (i.e. tool-complement, object complement or location-modifier). Body parts and natural kinds are expected to form the set of such entities.

Entity non-terminals
These are perceptible entities that consist of entity terminals in certain spatial configuration. They participate as complements or modifiers in more than one motor programme. Scenes are included in this set, and they participate in motor programmes as modifiers of actions (i.e. they denote the location in which an action takes place).

Having presented the basic action features, we can now turn to the definition of the 'vocabulary' of the action grammar:

Action grammar terminals
These are the simplest actions, i.e. perceptible movements carried out by an agent to achieve a goal, which have (one or more) body part tool-complements and no object complements. They have no action constituents themselves and they may be circular/repetitive. This is the set of all possible human body movements, such as limp, sprint, extend arm, raise hand, stretch leg, open/close hand, etc., i.e. the set of intransitive biological actions. Action terminals are further distinguished from each other through their perceptible motor features such as speed, force and direction. For example, the leaf movement nodes of the action parse tree in figure 1 comprise the 'extend' (hand$_1$) terminal and the 'enclose' (hand$_1$) terminal.

Action grammar non-terminals
These are perceptible action phrases that consist of action terminals (or other non-terminals) in a certain temporal configuration; they may have both tool-complements and object complements. They involve interaction with objects beyond one's own body or with other agents, for attaining a particular goal/task, such as grasp_knife, slice_tomato, etc. Searching for the value of an action phrase's complement using the

values of a subsequent action phrase complement guides the merging process (i.e. it determines the boundaries of the complex action, the top node). The latter implies that we define events as actions in temporal conjunction (sequential or parallel) that share features (see more details in the following section).

(c) Action grammar rules

Having defined the constituent elements of the action grammar, we can now present the production rules. These are presented in table 2.

In table 2, the action grammar production rules express the fact that no matter how simple or complex an action is, it has a compulsory goal specifier and a compulsory tool complement. The presence of affected object complements is optional and so is the presence of location modifiers. One will note that in this grammar, there is no explicit reference to the agent that performs the action, as in the language trees for example, where the actor (subject of the verb) may be denoted explicitly (corresponding noun phrase) or implicitly (through person agreement). This is so, because of the body-part tool complements of the action terminals; an action is inherently performed by an agent and since the terminals of the grammar incorporate necessarily a (human or animal) body-part complement, the agent information does not need to be explicitly present as a separate, non-tool complement.

Going back to the action tree in figure 1, the tree can be derived bottom-up, through recursive application of the grammar rules. Traditionally, parsers apply a grammar for the analysis of a certain structure, i.e. their input is the whole structure (e.g. sentence) to be analysed, segmented into tokens (terminals); when more than one sentence is to be parsed (i.e. a paragraph or whole text), automatic segmentation of the text into sentences is also provided in advance. In the language/symbolic space, the automatic segmentation of a text into sentences (i.e. structures to be considered separately for derivation of syntactic trees) and tokens (terminals) is a straightforward process owing to the discrete nature of the data to be analysed, whereas in the sensorimotor space, this is not the case[1]. Segmentation of a continuous stream of visual and/or motoric action into 'sentences' (i.e. groups of sequential or parallel actions that combine into an action tree) is a very challenging task [54]. Tokenization of such 'action sentences' requires a number of sensorimotor processing technologies to be employed, such as image segmentation, object recognition and action recognition; these technologies face a number of challenges and their output cannot be taken for granted when developing an action parser. On the contrary, a parser that applies the minimalist grammar of action can actually *use the grammar to guide the segmentation of visual and motoric action*:

— for 'tokenization': the parser can provide an 'attention-guiding' strategy for selecting those objects in a scene that are related to the actions as tools or affected objects (or even locations), while
— for 'sentence segmentation', the parser can use the minimalist action grammar operators and features

Table 2. Generative action grammar production rules. A, action terminal; A′, intermediate action structure; A″, maximal action structure; g, goal; m, modifier; t_c, tool complement; o_c, object complement; parentheses, optional presence.

rules	
4	$A'' \rightarrow g, A'$
3	$A' \rightarrow (m), A'$
2	$A' \rightarrow A', (o_c)$
1	$A' \rightarrow A, t_c$

to decide when merging stops and a new action structure starts.

In what follows, we sketch such a dynamic parser, which applies the rules of the minimalist grammar of action bottom-up along with related segmentation criteria in order to derive the parse tree(s) of non-previously segmented sensorimotor input.

Step 1:
— Find the first action A'_1 such that its start time is on/after the start time of the visual/motoric input A_{max}, and it has a body-part tool complement.

In other words, get the first body part that is in motion in the input stream A_{max} and keep its motoric characteristics as those pertaining to an action terminal A. This is a merging of an action terminal A with its effector, i.e. an entity that has a semantic type feature, the value of which can be unified with the tool-complement feature value of A, e.g. A = 'extend' [$t_{c\text{-}body\_part}$: Variable$_1$] merges with E″ = 'hand$_1$' [+ body part]. This creates the first binary action structure of the form: $A'_1 = \{A, E''\}$, e.g. A'_1 = 'extend hand$_1$'. Up to this point, *rule 1 of the action grammar* has been applied.

— Search for an entity that could satisfy an object complement feature of A'_1, i.e. for an object affected by the A'_1 action-tool binary structure (and no other action), and perform one more merging, creating the action structure A'_{1b}; if no such object complement is present, then a null object complement is derived.

This applies rule 2.

— If an object complement is present, get its location and create the action structure A'_{1c} which comprises A'_{1b} and the location as its modifier.

This applies *rule 3* of the grammar, only in cases when an object complement is present. At this stage, the parser does not proceed to applying rule 4, i.e. attributing a goal to the action structure; instead, it checks for what follows in order to decide whether to merge the following actions into a larger action structure.

Step 2:
— For as long as A_{max} extends in time beyond the end of the thus far derived structure A'_{1b} (i.e. as long as visual/motoric input is fed to the parser), find an action A'_2 that follows (or is parallel in time with) A'_{1b}, such that A'_{1b} and A'_2 share the same tool complement, or the tool complement of one is

the same with the object complement of the other; if so, relate the two actions through temporal conjunction and apply the grammar rules from the start.

For example, in the action parse tree in figure 1, the action 'enclose hand$_1$' (i.e. opening/closing hand for grasping) follows the 'extend hand$_1$' action; the fact that they share the same tool complement (i.e. hand$_1$) indicates that the two of them together form a more complex action unit. This is used as a criterion for continuing the merging of subsequent actions into the same action tree. Rule 1 is applied for the formation of A$_2'$ and rule 2 is applied too for interpreting all its complement features expanding it into A$_{2b}'$. So, a further merging takes place between the 'enclose hand$_1$' structure and the object that interprets its object complement feature, the 'knife'. Therefore, the A$_{2b}'$ action structure derived is 'enclose with hand$_1$ the knife'.

Since the merging of the subsequent actions has been decided, the parser performs another check:

— if A$_{1b}'$ had a null object complement, then attribute a 'reference' feature to this null complement and bind it with the A$_{2b}'$ object complement.

This is a binding between the object that the first action is directed to and the object that is indeed affected by the subsequent action; in other words, the direction of the first action of the sequence (e.g. 'extend hand') functions as a deictic reference to something that becomes obvious when the subsequent action is executed (i.e. the object 'knife' in our example—see reference feature in the parse tree in figure 1). Such decision regarding the deictic nature of an action can only be taken once the subsequent related action is found; i.e. the difference between 'extending a leg' and 'extending a leg towards X' is determined by the actions that follow these, such as 'extending another leg' (e.g. to stretch one's body) and 'kicking a ball', respectively.

The rules of the grammar may be applied recursively as actions combine in time-sharing complements; for example, the 'grasp with hand$_1$ knife' action shown in figure 1, may be followed by a 'pin with knife apple' action that extends the derived action structure further, adding one more constituent. This constituent is not just following the previous one in time, but its tool complement is shared with the object complement of the previous one. One can imagine infinite recursive applications of step 2 of our parser, e.g. adding one more constituent to our example, such as 'push with apple the plate' (an action whose tool complement is shared with the object complement of the previous one). Recursion in step 2 is guided through the correlation of the tool complement of subsequent actions; so, an action 'merges' with its tool (and optionally with an affected-object complement) and then 'merges' with a subsequent (or parallel) movement if they have the same tool complement (e.g. *extend hand—grasp with hand X*), if the object complement of the preceding one is the tool complement of the one that follows (e.g. *grasp with hand knife—cut with knife bread*) or vice versa.

— If the action that follows a thus far derived structure does not share a tool or object complement with the preceding one, a new action substructure of A$_{max}$ is created and processed applying the grammar rules bottom-up.

For example, the action sequence {extend hand$_1$—enclose with hand$_1$ knife} may be followed with {extend hand$_2$}. In this case, no clues exist that this action forms part of the thus far derived structure and therefore a new action substructure in A$_{max}$ is created (i.e. a second group of action constituents) and parsing starts again applying rule 1.

Step 2 is applied until no more actions are available in the input visual/motoric stream. In this sense, A$_{max}$ comprises an ordered set of action substructures, each substructure being an ordered set of actions itself.

In order for the parser to make the final decision regarding the merging of the action substructures, and thus, the derivation of one or more maximal action structures, two more steps are undertaken. These steps are the ones that lead to the application of rule 4 of the grammar, i.e. the attribution of the final goal to the maximal action structures.

Step 3:
— For each set A$_{simple}$ of constituents of A$_{max}$, which comprises only of the merging of an action terminal and its effector, find the first subsequent action-constituent set A$_{subsequent}$ whose first element F is not a simple action and shares the same tool complement with the A$_{simple}$ action. Expand A$_{simple}$ with a trace that is linked to F and is temporally combined to A$_{simple}$; the object complement of F is also bound with the null object complement of the A$_{simple}$ constituent through a reference feature sharing.

This step is applied in cases such as the one depicted in figure 2; there are two grasping actions (grasping a knife and grasping an apple) that take place in the following order: 'extend hand$_1$', 'extend hand$_2$', 'enclose with hand$_2$ apple', 'enclose with hand$_1$ knife' and so on. The action 'enclose with hand$_1$ knife' has a tool complement that is not shared with the preceding action, but it is shared with the first one. This is a case of a discontinuous action structure, a long-range dependency (see also section on action characteristics). This step of the action grammar implementation deals exactly with such phenomena, applying a *transformation*, on the derived action structure that allows a further *merging* of discontinuous actions (figure 3).

— For each set A$_{complete}$ of constituents of A$_{max}$, find the first subsequent action-constituent set A$_{subsequent}$ that comprises of at least one action structure with a tool complement common with an object complement of a constituent of A$_{complete}$. Link the two sets into the same complex action structure, though discontinuous in time.

This case of step 3 deals with disruption phenomena between more complex structures than the ones presented earlier (cf. also examples in §4).
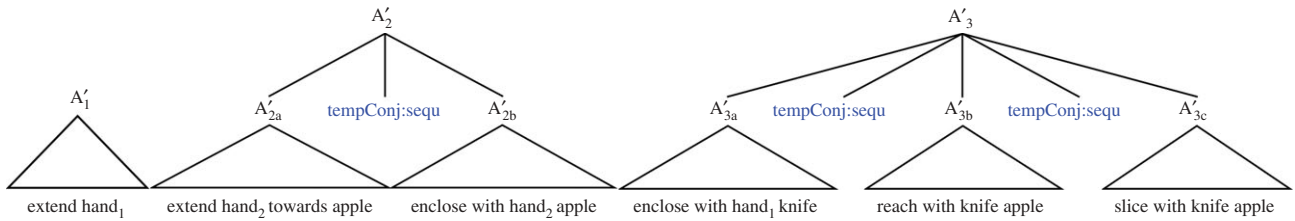
Figure 2. Incomplete parsing of the sequence: 'extend hand$_1$', 'extend hand$_2$', 'enclose with hand$_2$ apple', 'enclose with hand$_1$ knife', 'reach with knife apple' and 'slice with knife apple'. After step 2 of the action parser, three stand-alone action structures are derived rather than one structure comprising all three of them with the final goal of slicing.
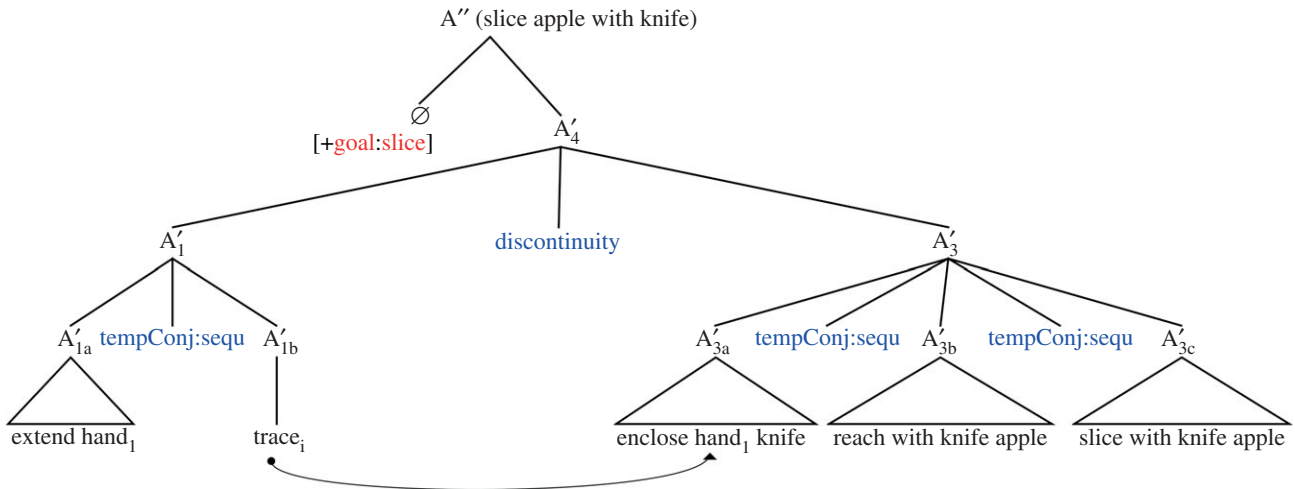

Figure 3. Action parse tree of the structure after applying step 3, i.e. after the move operation has been applied. A'_{3a} shares the same tool complement with A'_{1a} (hand$_1$) and its object complement (knife) is referred to by A'_{1a}. Therefore, its expected position is semantically exactly after A'_{1a} in position A'_{1b}. However, owing to the disruption by other actions, this position is empty; the action is in position A'_{3a}. Thus, a 'trace' of the action is left in position A'_{1b}, which is linked with the action in position A'_{3a}. The two structures A'_1 and A'_3 are not temporally combined, they are discontinuous; the actions that intervene in between may or may not be part of the same action structure. This is what step 4 checks.

Step 4:
— Conclude with merging subsequent constituents of A$_{max}$, that share one or more object complements.

This last step is needed (i) for deciding whether actions that intervene (in terms of time sequence) between linked discontinuous action structures, belong to the same maximal action structure (figure 4), and (ii) for unifying action structures that share object complements at any distance, as usually the case in events with a loose structure (loose in terms of temporal sequence of constituents and presence/absence of some constituents), e.g. *rinse tomato, grasp knife, bring bowl, cut tomato with knife, pour oil into bowl*, etc. for *preparing a salad*.

— When no more merging can take place, check the goal feature of each maximal action structure (i.e. apply rule 4 of the grammar) and exit.

Functions related to finding the tool of an action and the affected object are needed for the algorithm to work and are actually vital. As shown earlier, perceptual (e.g. visual) identification of body parts is a *sine qua non* requirement in this process and so is the notion of spatial intersection. The latter refers to a recursive 'merging' of body parts and objects as one comes into contact with another. So, the tool of an action is any object that is either a moving body part or a moving object spatially intersected with an effector and in synchrony with the effector. In the suggested algorithm, body parts and their intersection with other objects define not only the tools of an action, but also the affected objects, i.e. objects that are spatially intersected with a tool but they do not have the same motoric characteristics (e.g. one is not moving the other is, or they are both moving though not synchronized).

Note in the earlier-mentioned text that the tool complements of the action constituents of an action structure is an ordered set of entities that cannot be empty (it is defining for any action); a body part is the simplest tool, while its extension with other artefacts through a number of (tool-making) actions may form an infinitely complex tool.

Constraints on the use of body parts/effectors and their natural motors synergies should be incorporated in the algorithm sketched in this section, so that generation of correct and only correct action structures is guaranteed. The repertoire of possible motor synergies in human action should be taken into consideration. The suggested algorithm can deal with parallel syntax with slight modifications (e.g. checking not only for action constituents in a sequence, but in parallel timings too).
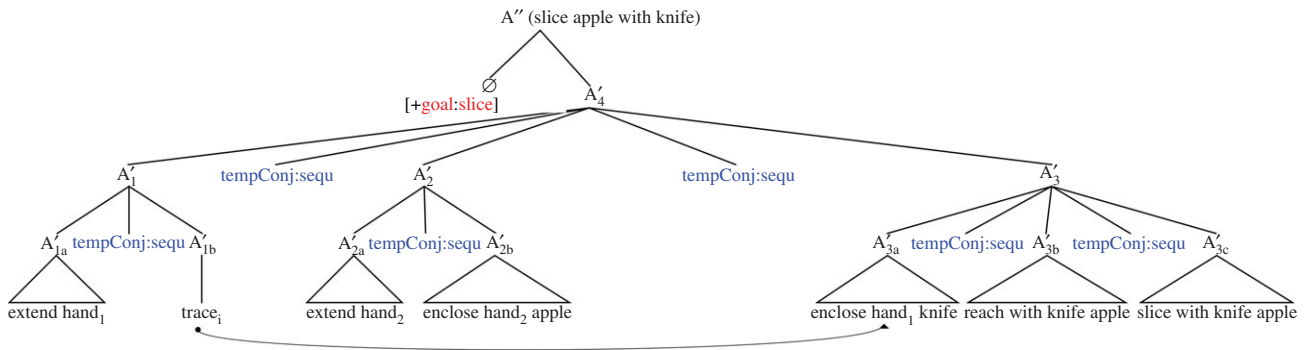
Figure 4. The maximal action structure for the sequence: 'extend $hand_1$, grasp with $hand_2$ apple, grasp with $hand_1$ knife, reach with knife apple, slice with knife apple'. After linking the discontinuous sub-action constituents in step 3 of the parser, a decision is taken regarding the action structure A'_2 that intervenes temporally causing the disruption: in applying step 4 of the parser, structure A'_2 is found to share an object complement with constituents of the action structure A'_3 (i.e. the apple). This is enough for considering A'_2 to be a constituent of the more complex action structure A'_4. Thus, the three independent action structures presented in figure 2, which were not combined into a common structure owing to a disruption phenomenon, are now all linked into a sequence with a common final goal.

The algorithm can be used beyond human action to animal action. For non-biological actions, the 'tool' of the action is any natural force exerted on an object (e.g. a door closing because of the wind); employing methodologies for identifying such forces (e.g. the use of language for describing what is going on in a video) can lead to use of the earlier-mentioned algorithm for analysing non-biological actions too.

4. ACTION GRAMMAR: RECURSION, MERGE, MOVE AND THE NOTION OF TOOL USE

We have used a minimalist framework for defining a generative grammar of action; however, is such a type of grammar really necessary? Would not a regular grammar or a context-free grammar be adequate for a formal analysis of the structure of action? In other words, is *recursion*, *merging* and *move*, all necessary for an action grammar?

Recursion is an important feature of generative grammars and has been shown to manifest itself in human language both:

— as tail recursion, a procedure that invokes another instance of itself as a final step, or in grammar terms, the embedding of a structure at the end of a structure of the same type [55], e.g. 'the man who knows your sister who works at the bookshop'. This is a complex noun phrase in which an anaphoric sentence ('who works...') is embedded at the end of another anaphoric sentence ('who knows...');
— as nested or true recursion, a procedure that invokes another instance of itself in mid-computation and then must resume the original procedure from where it left off, or in grammar terms, the embedding of a structure at the centre of a structure of the same type [55], e.g. 'the cat the boy saw left'. This is a sentence in which a sentence of the same type is embedded, interrupting its structure, and thus, creating a discontinuous structure with long-distance dependencies.

In the action grammar presented earlier, both types of recursion take place and they are both guided by the tool-use notion:

— tail recursion: this is recursion that takes place at step 2 of the algorithm presented above; it is the extension of an action with tool T and object O, with a following action with tool O on object X, which may be further extended with another action with tool X on object Y and so on; e.g.

'extend $hand_1$—grasp with $hand_1$ knife—cut with knife bread', or in language terms:
'extend $hand_1$, which grasps knife, which cuts bread'.

This is a complex action sequence, in which the third sub-action ('cut ...') is embedded at the end of the second sub-action ('grasp...'). The role of the tool used in an action structure is vital in determining the recursion.

— true recursion: this type of recursion may appear in action sequences, in cases when one starts doing something before finishing off with something else; e.g.

'extend $hand_1$—extend $hand_2$, grasp with $hand_2$ ball—grasp with $hand_1$ glass'.

This is an action sequence in which the 'grasp glass' action is interrupted in its execution by another action ('grasp ball') of the same type; this results in a discontinuous structure and creates a long-distance dependency between part of the 'grasp glass' action. Such true recursion may manifest at action structures of a varying degree of complexity, i.e. at complex events such as preparing a salad:

'grasp with $hand_1$ knife—grasp with $hand_2$ cutting board, press with cutting board cloth—cut with knife bread', or in a more complex level:
'grasp with $hand_1$ knife, pin with knife bread—grasp with $hand_2$ fork, pin with fork cheese, lick with tongue cheese—bite with teeth bread'.

In this case, the embedding disrupts the sequence of grasping a knife and actually using it, with an

Figure 5. Using techniques from machine learning, one can develop new nonlinear filters that—when applied to an image—produce a new image, where the intensity of a pixel is proportional to the probability that the pixel lies inside the image of a specific category, e.g. silverware. On the left is the image, and on the right the output of the filter.

action structure of the same type. Thinking of every-day tasks, such embeddings seem quite frequent; of course, given that in the motor space some actions can take place in parallel, such embeddings are not always found in a neat sequence with the discontinuous elements, but rather part of them overlaps in time. Considering interaction with other people for performing a task, the phenomenon becomes even more frequent; actions of one agent 'interrupt' those of another or, seen from another perspective, one agent compensates for missing needs for completing a task before the other agent asks for them (e.g. brings a cutting board for cutting the bread). It is a case of true coordination between actors.

Step 3 in the earlier-mentioned algorithm addresses such discontinuous cases. Again, the notion of tool use is employed for guiding the combination of discontinuous elements; it is the tool complement of different actions that binds them together, though discontinuous in time.

Recursive *merging* in human language initiates a search for elements that solve variables in morpho-syntactic features of words/phrases; this is fundamental in the suggested action grammar too, because it guides all derivation. Through this search and merge mechanism, actions combine with tools and affected objects and with other actions through unification of their features. Going beyond the merging of actions and corresponding tools/objects, merging in the action grammar takes place between elements of the same type too; this is what has been called in the minimalist framework 'internal merging' or 'move' [47]. Compare with, for example, the 'extend hand to X' sub-action in figure 1, in which there is clearly a reference to an object complement. The actual object complement forms part of a subsequent action. This is a probe–goal relation between the object complement feature of one action and the corresponding feature of a subsequent one. Feature binding is common in action (owing to object permanence); however, the reference mechanism manifests itself only in those cases that involve directed motion towards objects, but no contact with them. Furthermore, all 'disruption' cases (true recursion) mentioned earlier are 'internal merging' cases too. Step 3 of the parsing algorithm makes use of the move operation, leaving a trace at the position the constituent should normally be found and linking the trace with the constituent in its actual position in the action structure.

On the basis of all the earlier-mentioned accounts, we suggest that the generative grammar of action must necessarily allow for both tail and true recursion, and the use of merging and move operators. In such grammar, tool use plays an important role for the derivation of action structures.

5. DISCUSSION

In employing a generative grammar for describing the structure of action, one substantiates experimental evidence on the common biological basis of language and action, and feeds the long-standing debate on language evolution and what makes human language unique [12,56,57]. However, why would one need a grammar of action and what does it mean, if tool use is indeed the computational structuring principle of action?

As shown already, a grammar of action has been sought for in both artificial intelligence and cognitive science; generalization, learning and prediction of action in both human cognition and computation depends on identifying a structure of action that guides action-related processing in both action recognition and action generation. Starting with artificial intelligence, event recognition and visual scene understanding have been the applications most interested in identifying a perceptual grammar of action for endowing robots and machines with the skills to recognize and interpret human behaviour. Large-scale video-processing depends on robust tools that perform visual object and visual action recognition; according to the suggested action grammar, recognition of human body parts ([58]; figures 4 and 5) is the key to such applications and drives action recognition and in particular motor primitive recognition. Recognition of the spatial intersection of body parts and other objects is the next most important tool needed; this is technology that segments objects robustly ([59]; figures 6–9) going beyond the visual merging of objects (e.g. the extension of a body part with the grasped object) and identifying not necessarily the type of object but instead its role as tool or object of interaction according to its spatial relation to a body part or an extended body part. On the basis of these two technologies, the action grammar can be used

Figure 6. Just like we can learn filters for objects, we can also learn them for body parts, legs, arms, heads, torsos and hands. On the left is the image; on the right is the output of filters for body parts denoted in different colours.
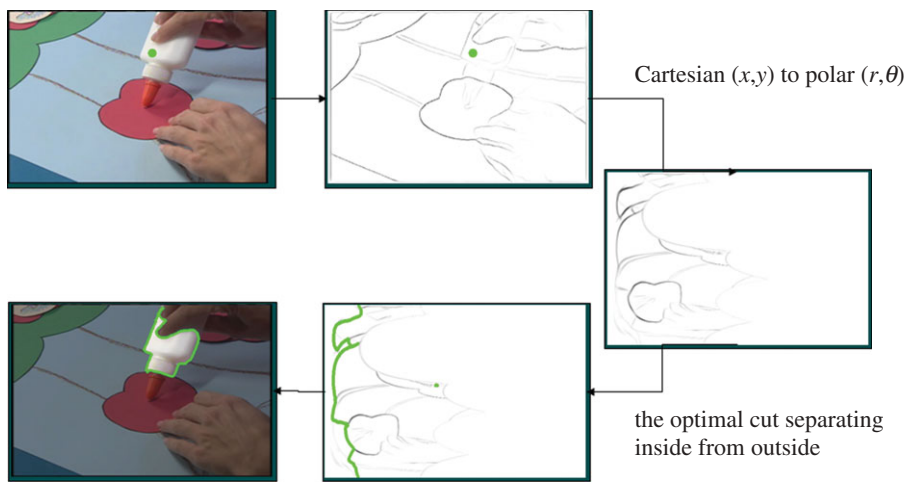
Figure 7. By fixating at a part of a scene (selecting a point in the image), we can segment the object containing the fixation point, in this case a glue bottle from a hands and crafts activity. Images and video courtesy of Johns Hopkins Workshop on Vision and Language.
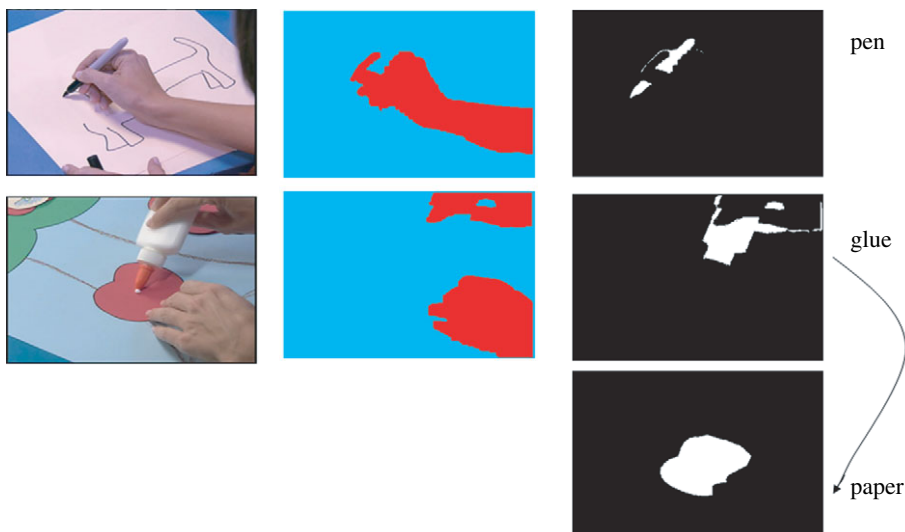
Figure 8. By using the filters described in figures 5 and 6, we can process videos of human activity to segments hands, tools and objects participating in actions. Top row: left: image from a 'drawing' activity; middle: segmentation of hands; right: segmentation of the object in the hand (pen). Bottom row shows results from another activity. Images and video courtesy of Johns Hopkins Workshop on Vision and Language.

| image of activity | scissors and hand filter | hand segmentation | scissors segmentation |
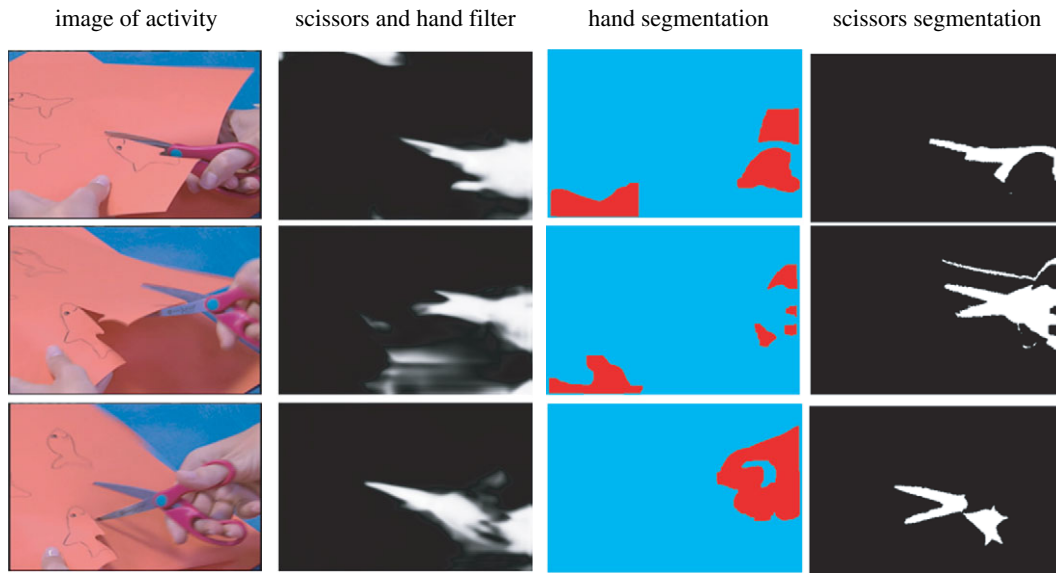
Figure 9. Image of the activity (cutting paper with scissors), scissors and hand filter output, hand segmentation, scissors segmentation using the technique of figure 5 and the filter output. Images and video courtesy of Johns Hopkins Workshop on Vision and Language.

for parsing actions of any complexity, without ever going into full identification of the objects involved in these actions. All recognition is based on the pragmatic roles of the objects.

Going to action generation, robotics is interested in advanced motor control that allows an agent to plan the execution of an action (global control strategies) by combining motor primitives into actions that lead to attaining the final goal. In this task, putting actions in sequence and coordinating the use of the robot's effectors linearly or parallel in time for achieving a task is usually hard-coded and strictly dependent on the exact action that is to be executed. The suggested action grammar can be used as a sophisticated motor control planner that will generate correct and only correct sequences of actions depending on objects that the robot sees in its environment, experimenting with the different roles (tool or object of interaction) to be assigned to each object and with the execution of motor primitives; this is a guided object manipulation and exploration that can be used as a method by the robot to learn new behaviours, without necessarily being able to identify the exact type of objects. The grammar provides a way to determine the endpoint of a sequence of actions, without relying on knowledge of the exact action type.

As shown in the previous sections, the minimalist grammar of action comprises of features whose importance in an action grammar is corroborated by neurobiological evidence and so is the hierarchical and compositional nature of action structure. Action structures in our grammar are derived through *merging* which is a very basic operation, that of composition, and so is the move operation (since it is defined as 'internal merging'). However, *what drives* the merging in our action grammar, calls for thorough exploration through experimentation. In the human action space, this question is in many ways equal to the question of what actually drives attention. In the minimalist framework, it is the features that drive the merging. In

our action grammar, it is indeed the tool and affected object complements that drive all merging, with the former playing a major role in all derivations.

In human cognition, there is a growing literature on the importance of the notion of tool use [13,60]. However, no experiments have been reported on the role of this notion for structuring action. If our argument that 'tool use' as the structuring principle of action has a biological basis (rather than merely a computational one), one would expect that an inability in humans to attribute the 'tool' role to an object within an action would be associated with inability to recognize or produce the hierarchical, recursive structure of an action.

Closer to the envisaged experiments are ones that show aphasics having problems in sequencing biological actions (e.g. to serve a cup of tea), while they have no problem in sequencing non-biological events (e.g. a bicycle falling) [14]. In this study, patients were also found to have severe problems in naming tools and tool use, while they understood the global meaning of what they had seen. In a follow up of this work, it has been found that it is the ordering of *transitive and 'syntactic' biological actions* in particular that is affected by virtual lesions in the left Broca's area 44 [61]. These actions involve hand–object interaction (i.e. tool–object interaction, e.g. cutting something, as opposed to non-transitive ones such as 'getting up'), and have a compositional structure; they correspond to the ones that are derived through recursion in our action grammar.

So is it the attribution of the tool role or the mechanism of recursion that is affected in such cases, or even both? Tool use and language have been claimed to share computational mechanisms for processing complex hierarchical structures [60], a capacity that exists in primates with no language (of the complexity of human language) and that could have been exapted to support human grammatical ability [62]. Tool making in particular has been speculated to have provided to action representation the capacity of recursion [13].

Through the minimalist grammar of action, we argue that action structure is recursive and it is tool use that drives both merging (including move) and recursion.

ENDNOTE

[1] We refer to tokenization in text, on which parsers normally run; in speech, the tokenization difficulties owing to the 'continuous' nature of the data processed become more evident.

REFERENCES

1 Lashley, K. 1951 The problem of serial order in behaviour. In *Cerebral mechanisms in behaviour* (ed. L. Jefress), pp. 112–137. New York, NY: Wiley.

2 Leroi-Gourhan, A. 1964 *Le geste et la parole*, vols. 2. Paris, France: Albin Michel.

3 Bauer, P. 1995 Recalling past events: from infancy to early childhood. *Ann. Child Dev.* **11**, 25–71.

4 Whiten, A., Flynn, E., Brown, K. & Lee, T. 2006 Imitation of hierarchical action structure by young children. *Dev. Sci.* **9**, 574–582. (doi:10.1111/j.1467-7687.2006.00535.x)

5 Allen, K., Ibara, S., Seymour, A., Cordova, N. & Botvinick, M. 2010 Abstract structural representations of goal-directed behavior. *Psychol. Sci. J. Am. Psychol. Soc.* **21**, 1518–1524. (doi:10.1177/0956797610383434)

6 Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. 2000 Visuomotor neurons: ambiguity of the discharge or 'motor' perception? *Int. J. Psychophysiol.* **35**, 165–177. (doi:10.1016/S0167-8760(99)00051-3)

7 Fogassi, L., Ferrari, P., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. 2005 Parietal lobe: from action organization to intention understanding. *Science* **308**, 662–667. (doi:10.1126/science.1106138)

8 Aloimonos, Y., Guerra-Filho, G. & Ogale, A. 2009 The language of action: a new tool for human-centric interfaces. In *Human centric interfaces for ambient intelligence* (eds H. Aghajan, J. Augusto & R. Delgado), pp. 95–131. New York, NY: Academic Press.

9 Aloimonos, Y. 2008 HAL: human activity language. *J. Vision* **8**. (doi:10.1167/8.6.1050)

10 Guerra-Filho, G. 2007 A sensory-motor linguistic framework for human activity understanding. PhD thesis. Department of Computer Science, University of Maryland, College Park.

11 Chomsky, N. 1995 *The minimalist program*. Cambridge, MA: MIT Press.

12 Chomsky, N. 2005 Three factors in language design. *Linguist. Inq.* **36**, 1–22. (doi:10.1162/0024389052993655)

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

14 Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A., Pozzo, T., Calzolari, F., Granieri, E. & Fadiga, L. 2009 Encoding of human action in Broca's area. *Brain* **132**, 1980–1988. (doi:10.1093/brain/awp118)

15 Flash, T. & Hochner, B. 2005 Motor primitives in vertebrates and invertebrates. *Curr. Opin. Neurobiol.* **15**, 660–666. (doi:10.1016/j.conb.2005.10.011)

16 Viviani, P. 1986 Do units of motor action really exist? In *Generation and modulation of action patterns* (eds H. Heuer & C. Fromm), pp. 201–216. New York, NY: Springer-Verlag.

17 Mussa-Ivaldi, F. & Bizzi, E. 2009 Motor learning through the combination of primitives. *Phil. Trans. R. Soc. B* **355**, 1755–1769. (doi:10.1098/rstb.2000.0733)

18 Hart, C. & Giszter, S. 2004 Modular premotor drives and unit bursts as primitives for frog motor behaviours. *J. Neurosci.* **24**, 5269–5282. (doi:10.1523/JNEUROSCI.5626-03.2004)

19 Stein, P. 2005 Neuronal control of turtle hind limb motor rhythms. *J. Comput. Physiol.* **191**, 213–229. (doi:10.1007/s00359-004-0568-6)

20 Roitman, A., Massaquoi, S., Takahashi, K. & Ebner, T. 2004 Kinematic analysis of manual tracking in monkeys: characterization of movement intermittencies during a circular tracking task. *J. Neurophysiol.* **91**, 901–911. (doi:10.1152/jn.00261.2003)

21 Pasalar, S., Roitman, A. & Ebner, T. 2005 Effects of speeds and force fields on submovements during circular manual tracking in humans. *Exp. Brain Res.* **163**, 214–225. (doi:10.1007/s00221-004-2169-6)

22 Fishbach, A., Roy, S., Bastianen, C., Miller, L. & Houk, J. 2005 Kinematic properties of on-line error corrections in the monkey. *Exp. Brain Res.* **164**, 442–457. (doi:10.1007/s00221-005-2264-3)

23 Jeannerod, M. 1994 Object oriented action. In *Insights into the reach and grasp movement* (eds K. Bennett & U. Castiello), pp. 3–15. London, UK: Elsevier Science.

24 d'Avella, A., Saltiel, P. & Bizzi, E. 2003 Combinations of muscle synergies in the construction of a natural motor behaviour. *Nat. Neurosci.* **6**, 300–308. (doi:10.1038/nn1010)

25 Tresch, M., Saltiel, P. & Bizzi, E. 1999 The construction of movement by the spinal cord. *Nat. Neurosci.* **2**, 162–167. (doi:10.1038/5721)

26 Cheung, V., d'Avella, A., Tresch, M. & Bizzi, E. 2005 Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviours. *J. Neurosci.* **25**, 6419–6434. (doi:10.1523/JNEUROSCI.4904-04.2005)

27 Ting, L. & MacPherson, J. 2005 A limited set of muscle synergies for force control during a postural task. *J. Neurophysiol.* **93**, 609–613. (doi:10.1152/jn.00681.2004)

28 Ivanenko, Y., Cappellini, G., Dominici, N., Poppele, R. & Lacquaniti, F. 2005 Coordination of locomotion with voluntary movements in humans. *J. Neurosci.* **25**, 7238–7253. (doi:10.1523/JNEUROSCI.1327-05.2005)

29 Arbib, M. 1992 Schema theory. In *The Encyclopedia of artificial intelligence* (ed. S. Shapiro), pp. 1427–1443. New York, NY: Wiley Interscience.

30 Jeannerod, M., Arbib, M., Rizzolatti, G. & Sakata, H. 1995 Grasping objects. The cortical mechanisms of visuomotor transformation. *Trends Neurosci.* **18**, 314–320. (doi:10.1016/0166-2236(95)93921-J)

31 Juhola, M. 1995 A syntactic analysis method for eye movements of vestibule-ocular reflex. *Comput. Methods Programs Biomed.* **46**, 59–65. (doi:10.1016/0169-2607(94)01599-B)

32 del Vecchio, D., Murray, R. & Perona, P. 2003 Decomposition of human motion into dynamics-based primitives with application to drawing tasks. *Automatica* **39**, 2085–2098. (doi:10.1016/S0005-1098(03)00250-4)

33 Jenkins, O. & Mataric, M. 2003 Automated derivation of behavior vocabularies for autonomous humanoid motion. In *Proc. Int. Conf. Autonomous Agents Multi-Agent Systems, Melbourne, Australia, July*, pp. 225–232. (doi:10.1145/860575.860612)

34 Rohrer, B., Fasoli, S., Krebs, H., Hughes, R., Volpe, B., Frontera, W. R., Stein, J. & Hogan, N. 2002 Movement smoothness changes during stroke recovery. *J. Neurosci.* **22**, 8297–8304.

35 Grinyagin, I., Biryukova, E. & Maier, M. 2005 Kinematic and dynamic synergies of human precision-grip movements. *J. Neurophysiol.* **94**, 2284–2294. (doi:10.1152/jn.01310.2004)

36 Kang, N., Shinohara, M., Zatsiorsky, V. & Latash, M. 2004 Learning multifinger synergies: an uncontrolled manifold analysis. *Exp. Brain Res.* **157**, 336–350. (doi:10.1007/s00221-004-1850-0)

37 Jerde, T. & Flanders, M. 2003 Coarticulation in fluent fingerspelling. *J. Neurosci.* **23**, 2383–2393.

38 Fod, A., Mataric, M. & Jenkins, O. 2002 Automated derivation of primitives for movement classification. *Auton. Robots* **12**, 39–54. (doi:10.1023/A:1013254724861)

39 Kahol, K., Tripathi, P. & Panchanathan, S. 2004 Automated gesture segmentation from dance sequences. *Proc. IEEE Int. Conf. Autom. Face Gesture Recognit.* 883–888. (doi:10.1109/AFGR.2004.1301645)

40 Nakazawa, A., Nakaoka, S., Ikeuchi, K. & Yokoi, K. 2002 Imitating human dance motions through motion structure analysis. *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.* **2**, 2539–2544. (doi:10.1109/IRDS.2002.1041652)

41 Wang, T., Shum, H., Xu, Y. & Zheng, N. 2001 Unsupervised analysis of human gestures. *Proc. IEEE Pacific Rim Conf. Multimedia* **2195**, 174–181.

42 Ogale, A., Karapurkar, A. & Aloimonos, Y. 2007 View invariant modeling and recognition of human action using grammar. *Lect. Notes Comput. Sci.* **4358**, 115–126. (doi:10.1007/978-3-540-70932-9_9)

43 Chomsky, N. 1956 Three models for the description of language. *IRE Trans. Inf. Theory* **2**, 113–124. (doi:10.1109/TIT.1956.1056813)

44 Chomsky, N. 1957 *Syntactic structures.* Berlin, Germany: Mouton de Gruyter.

45 Chomsky, N. 1965 *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

46 Chomsky, N. 1993 *Lectures on government and binding: the Pisa lectures.* Berlin, Germany: Mouton de Gruyter.

47 Lasnik, H. 2002 The minimalist program in syntax. *Trends Cogn. Sci.* **6**, 432–437. (doi:10.1016/S1364-6613(02)01977-0)

48 Iriki, A. & Sakura, O. 2008 The neuroscience of primate intellectual evolution: natural selection and passive and intentional niche construction. *Phil. Trans. R. Soc. B* **363**, 2229–2241. (doi:10.1098/rstb.2008.2274)

49 Iriki, A., Tanaka, M. & Iwamura, Y. 1996 Coding of modified body schema during tool use by macaque postcentralneurones. *Neuroreport* **14**, 2325–2330.

50 Mantovani, G., Bufalari, I., d'Ausilio, A. & Fadiga, L. 2011 *The brain representation of objects and tools.* POETICON Project, D3.2b, Athens, Greece.

51 Luria, A. 1973 *The working brain.* London, UK: Penguin.

52 Cangelosi, A. *et al.* 2010 Integration of action and language knowledge: a roadmap for developmental robotics. *IEEE Trans. Auton. Mental Dev.* **2**, 167–195. (doi:10.1109/TAMD.2010.2053034)

53 Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G. & Rizzolatti, G. 2007 Impairment of action chains in autism and its possible role in intention understanding. *Proc. Natl Acad. Sci. USA* **104**, 17 825–17 830. (doi:10.1073/pnas.0706273104)

54 Li, Y. & Aloimonos, Y. 2009 The action synergies: building blocks for understanding human behavior. In *Proc. IEEE Int. workshop on social signal processing, Amsterdam, The Netherlands*, pp. 1–7. (doi:10.1109/ACII.2009.5349506)

55 Pinker, S. & Jackendoff, R. 2005 The faculty of language: What's special about it? *Cognition* **95**, 201–236. (doi:10.1016/j.cognition.2004.08.004)

56 Hauser, M., Chomsky, N. & Fitch, W. 2002 The faculty of language: What is it, who has it, and how did it evolve? *Science* **298**, 1569–1579. (doi:10.1126/science.298.5598.1569)

57 Fitch, W., Hauser, M. & Chomsky, N. 2005 The evolution of the language faculty: clarifications and implications. *Cognition* **97**, 179–210. (doi:10.1016/j.cognition.2005.02.005)

58 Summerstay, D. & Aloimonos, Y. 2010 Learning to recognize objects in images using anisotropic nonparametric kernels. In *Proc. 1st Annu. Meet. Biologically Inspired Cognitive Architectures (BICA) Society* (eds A. V. Samsonovich, K. R. Jóhannsdóttir, A. Chella & B. Goertzel), pp. 163–168. Amsterdam, The Netherlands: IOS Press. (doi:10.3233/978-1-60750-661-4-163)

59 Mishra, A. & Aloimonos, Y. 2009 Active segmentation. *Int. J. Humanoid Robot.* **6**, 361–386. (doi:10.1142/S0219843609001784)

60 Stout, D., Toth, N., Schick, K. & Chaminade, T. 2008 Neural correlates of early stone age toolmaking: technology, language and cognition in human evolution. *Phil. Trans. R. Soc. B* **363**, 1939–1949. (doi:10.1098/rstb.2008.0001)

61 Clerget, E., Winderickx, A., Fadiga, L. & Olivier, E. 2009 Role of Broca's area in encoding sequential human actions: a virtual lesion study. *Neuroreport* **20**, 1496–1499. (doi:10.1097/WNR.0b013e3283329be8)

62 Higuchi, S., Chaminade, T., Imamizu, H. & Kawato, M. 2007 Shared neural correlates for language and tool use in Broca's area. *Neuroreport* **20**, 1376–1381. (doi:10.1097/WNR.0b013e3283315570)

Review

The origins of non-human primates' manual gestures

Katja Liebal[1,2,3,]* and Josep Call[3]

[1]*Department of Education and Psychology, Cluster Languages of Emotion, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany*
[2]*Department of Psychology, University of Portsmouth, King Henry 1st Street, Portsmouth PO1 2DY, UK*
[3]*Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany*

The increasing body of research into human and non-human primates' gestural communication reflects the interest in a comparative approach to human communication, particularly possible scenarios of language evolution. One of the central challenges of this field of research is to identify appropriate criteria to differentiate a gesture from other non-communicative actions. After an introduction to the criteria currently used to define non-human primates' gestures and an overview of ongoing research, we discuss different pathways of how manual actions are transformed into manual gestures in both phylogeny and ontogeny. Currently, the relationship between actions and gestures is not only investigated on a behavioural, but also on a neural level. Here, we focus on recent evidence concerning the differential laterality of manual actions and gestures in apes in the framework of a functional asymmetry of the brain for both hand use and language.

Keywords: gesture; manual; ontogenetic ritualization; phylogenetic ritualization; laterality; great apes

1. INTRODUCTION

Dolphins are one of the most gracile and elegant creatures of the sea. However, before dolphins became what they are today, they underwent a remarkable transformation. The terrestrial ancestor of dolphins was a hippopotamus-like creature that walked on all fours and lacked the stylized forms, and presumably the elegant movements, of its marine descendant. Over the last 50 million years, dolphins have been evolving into what they are today. This remarkable transformation teaches us an important lesson. Complex structures such as legs and snouts can be transmuted over time into equally complex and functionally equivalent structures such as fins and blowholes, respectively.

The relation that exists between gesture and action is in some ways analogous to the relation that exists between fins and legs or between noses and blowholes. A central thesis of this contribution is that many of the gestures displayed by apes began their existence as actions devoid of a communicative function, but over time they became co-opted and transformed into communicative devices that accomplished similar functions [1]. Moreover, just like fins and legs, this change took place over evolutionary time, but in the case of gestures, it can also take place during the lifetime of one individual. In both cases, however, one can find clues that inform us about their origin. Comparing the communicative repertoires of monkeys and apes with those of humans can play a crucial role in the quest for the roots of human language, and more specifically in the role that gestures might have played in the evolution of language.

The focus of this paper is twofold. First, we will explore the question of how actions are transformed into gestures both from a phylogenetic and an ontogenetic perspective. To this end, we will present the latest advances in ape gestural communication, including some of the controversies in the field. We will begin by defining gestures, briefly presenting some of the features of the apes' gestural repertoires and discussing three ways in which individuals can acquire gestures. Second, having established the connection between actions and gestures, we will turn our attention to the role that gestures may have played in language evolution. First, we will note the increasing interest in gestural communication of our closest living primate relatives in the framework of the proposed close link between action and language in humans. Then we will link recent data on ape laterality in gestural use with language hemispheric specialization.

2. GESTURE ORIGINS (OUT OF ACTIONS)

(a) *Defining a gesture*

Human gestures are usually very broadly referred to as the 'manner of carrying the body' and 'movements of the body or limbs as an expression of feeling' ([2],

---

* Author for correspondence (katja.liebal@fu-berlin.de).

p. 476). According to Kendon [3], a gesture is a form of non-verbal communication in which visible bodily actions communicate particular messages, either in place of speech or together and in parallel with spoken words. Before children start to speak, they use a variety of gestures to communicate with their carers, such as showing objects and pointing to objects, events or persons in their environment [4–7]. Even when they acquire their first words, gestures are not simply replaced but are incorporated in their verbal communication [8,9].

In adults, a substantial body of research addresses the kinds of manual gestures produced by humans while speaking [3,10–12]. If spoken language is not possible, then manual gestures can convey very specific and complex information, even replacing spoken language, thus becoming a form of a highly conventionalized sign system [13,14]. Thus, in humans, gestures can vary in their degree of conventionalization and therefore the degree to which they are linked to or even replace spoken language ([15], pp. 37–40). Therefore, research into human gestures is a highly diverse field, since it covers very different kinds of gestural communication, such as speech-accompanying gestures, gestures of pre-linguistic children or even gestures co-occurring with sign language. The question arising here is whether non-human primate species, which are clearly lacking spoken language but with bodies and particularly fore-limbs sharing many characteristics with human beings, use gestures to communicate with conspecifics. To tackle this question and to enable any comparison with humans at all, we need to focus on human gestures not used in combination with language (either spoken or signed) and thus on the gestural communication of pre-linguistic children. By adopting the corresponding criteria, a gesture is defined as a behaviour that unlike an action is motorically ineffective. It requires the active participation of a partner to fulfil its purpose, it is produced in the presence of an audience and is tailored to the attentional state of the audience. Furthermore, it involves gaze alternation or visual checking between social partners and distant objects or events, is characterized by the sender's waiting for the recipient's response and displays persistence and elaboration of communicative behaviour when communicative attempts fail [16–19].

As our previous introduction to the term *gesture* pointed out, gestures are not restricted to the use of hands, but often include movements of limbs and also head and body movements, as well as postures. Some scholars even include facial expressions as gestures [20,21]. However, here we only focus on manual gestures in non-human primates, that is, gestures produced with the whole arm or hands. We also mainly discuss studies of gestural communication in great apes; this is not to neglect gibbons and monkeys, but so far there is still little evidence of hand use for the purpose of communication in non-great ape species ([22], but see [23–25]).

One of the biggest challenges in gestural research lies in determining when an instrumental action has crossed the threshold and becomes a gesture. Some gestures are easy to distinguish from instrumental actions, but there are others that are much more difficult to differentiate. For instance, we would include as

gestures the subtle touches and presses that dancers use to inform their partner about their impending actions or to direct them in a certain way. In contrast, we would not consider as gestures holding an infant when she is beginning to walk because here the main function would be to help the infant to maintain her equilibrium.

The problem of deciding between gestures and actions is further compounded when multiple species are considered. Although the potential for confusing actions and gestures represents a potential analytic weakness, it can become a strength since it tells us something about the origin of gestures. In particular, it suggests that at least some gestures may have begun their existence as actions before they were transformed into a communicative function.

From a more practical point of view, one approach that we find useful in distinguishing actions from gestures is to consider how many of the criteria outlined above are met. Thus, faced with a potential candidate as a gesture, we must ask whether (i) it is motorically ineffective, (ii) there is response waiting, (iii) gaze alternation, and (iv) persistence. The more criteria are met, the more sure we can be that a given behaviour qualifies as a gesture. One cannot be 100 per cent sure but at least this method can help in reducing our uncertainty.

In the next two sections, we present a brief overview of the gestural communication of the great apes (see [18] for a more detailed treatment) and then discuss their potential origins.

(b) Gestures of non-human primates

Unlike research into human gestures mostly restricting the focus on the visual channel [3], researchers investigating primate gestures also consider tactile gestures such as *push* or *throw objects* and gestures with an auditory component such as *hand clap* and *chest beat*. Including gestures that transfer information via non-visual channels captures the richness and subtleties of non-vocal communication. However, it also raises potential problems when it comes to distinguishing gestures from instrumental actions. For instance, a gesture called *reach* that consists of extending an arm in the direction of a conspecific is easier to identify as gesture than a gesture called *touch-side* that consists of touching an individual on her side to make her move. The reason for this is simple. The lack of physical contact between the two interacting individuals automatically makes *reach* motorically ineffective, one of the first criteria to identify a gesture as such. After all, it is conceivable that the *touch-side* gesture involved enough force to make the individual move, thus making this action motorically effective and automatically disqualifying it as a gesture. In sum, researchers investigating non-human primates have faced a trade-off between capturing the richness and subtleties of non-vocal communication in primates at the expense of making the distinction between gestures and instrumental actions less clear-cut than in human research.

In a recent summary of a systematic comparison of the four great apes, siamangs and Barbary macaques, Call & Tomasello [18] concluded that those species

differed in repertoire size, composition and function of their gestures. They reported between 20 and 35 different gesture types depending on the species, which meet the above-mentioned criteria of being motorically ineffective and are accompanied by response waiting and/or gaze alternation, as well as persistence in case the recipient did not react. Out of those reported gestural repertoires, at least 50 per cent of each species' repertoire consisted of manual gestures, with the highest proportion found in gorillas (73%). (It is important to note that those numbers refer to the total repertoires found across different groups of one species, not average proportions.) For example, tactile gestures, which included some kind of physical contact with another individual (e.g. *touch*, *pull* or *slap*), were used by all great apes, siamangs and Barbary macaques [18]. Auditory gestures often included the individual's own body used to produce that noise, such as *hand clap* in chimpanzees [26], and *chest beat* and *body beat* in gorillas [27]. Alternatively, noise can be produced by using objects while performing gestures such as *ground slap*, *push objects* or *throw objects*, which are gestures particularly reported for chimpanzees [26]. On the other hand, examples for silent gestures not involving physical contact include gestures like *extend arm* (*reach*), *arm raise* and *wave arm* [18]. As opposed to bonobos, siamangs and Barbary macaques, chimpanzees and orangutans often incorporated objects in their gestural displays (15% of the gestures). For example, orangutans offer food to other individuals by extending one arm with food in their hand to another individual [28] and chimpanzees use branches, which they shake vigorously to get the attention of another group member [26]. The higher values for gestures involving objects for chimpanzees and orangutans are interesting in light of their higher propensity to use tools in the wild than the other species and may be indicative of a common neural substrate for tool use and gestural communication.

So far, we have mostly presented the results of our own research project on gestural communication of non-human primates that started with the work by Tomasello *et al.* [29]. Of course, there are many more scholars working on the question of which gestures non-human primates use, how they acquire them and what the underlying socio-cognitive skills are, both in wild and captive settings. The first pioneering field studies report several gestures as parts of ethograms for orangutans [30], gorillas [31], chimpanzees [32] and bonobos [33], but also for gibbons [34,35] and monkeys [36]. Lately, there is an increase in more systematic, mostly observational studies investigating gesture use within social groups of great apes [37,38] and monkeys [22,24,39–41]. This increasing body of research reflects the interest in the role gestures might have played for the evolution of human language [42–45], although studies addressing facial expressions or vocalizations still outnumber studies concerning gestures [46].

However, the reported gestural repertoires for the different species vary considerably between studies. For instance, while Pika *et al.* [27] described 33 gestures for gorillas, Genty *et al.* [37] reported more than 100 gestures for this species. Furthermore, very different names are often used to categorize the same behaviour, complicating comparisons across studies and species. These discrepancies may be attributed in part to the sampling effort and the differences in the detail of the coding schema across studies [47], but it remains a fact that gestures are, first of all, difficult to differentiate from actions, and second, although the majority of gestures are not gradual signals like in the case of facial expressions, they are difficult to categorize because of the often different criteria used across studies to define a gesture. This is closely related to a third reason, namely that gestures are often defined based on their function or the context they are used in (e.g. *food offer*), resulting in a conflation of form and function rather than referring to form and meaning as separate variables.

However, although the paucity of data in terms of the number of species and groups investigated and also in terms of consistency of definitions used across studies prevents us from concluding that there are any systematic differences between species (yet), at the very least we can say that hands play an important role in gestural communication among primates.

Considering the function of gestural communication, monkeys and apes use the majority of their gestures to request actions like grooming, play or mating. They use their gestures in a dyadic way and usually not to communicate about events or objects outside their dyad, but to request certain actions, expecting an immediate response [48]. In case the recipient is not reacting, they will continue to gesture until they finally receive the appropriate response of their social partner [49–51]. Apes do take into account the visual access of others (see [52] for a review), use visual gestures only if the recipient is attending [18,53], or use other strategies like moving into someone's visual field before starting to gesture [50,54]. In other words, they take into account the behaviour of others and adjust their communicative means accordingly. However, there are inconsistent results as to to what extent apes are actually able to alter their gestures if their first gesture was not successful—for chimpanzees and orangutans, it is shown that most often the same gesture is repeated [50,55], while gorillas seem to show more flexibility in alternating the gestures they use to achieve a certain goal [49]. It should be considered that for interactions with a human experimenter, both chimpanzees and orangutans were shown to not only substitute, but also elaborate their gestures depending on the behaviour of a human in case their goal was not met [19,56].

Another much-debated topic is the question of pointing in non-human great apes. In captivity, great apes and also some monkey species point to request food, tools or particular actions from humans [57–61]. Pointing in great apes represents a flexible, intentional behaviour, since the use of this gesture is adjusted to the attentional state of the human and it occurs in combination with other signals such as facial expressions and vocalizations [19,57,62,63]. Pointing is also frequently used by language-trained apes [60,64,65], where it often resembles the form of the pointing gesture of Western cultures with the arm and index finger extended [66].

However, unlike human infants that also point to show objects, to share attention upon things, or even inform others [67], non-human primates usually point to request things or actions in their interactions with humans. The vast majority of great apes' pointing gestures therefore fall within the category of so-called imperative gestures, which consist of the ape using the gesture to obtain something that they want from the human (see [68] for a review). Great apes also inform a human by pointing to the location of a hidden tool, but with the aim that the human can use it to retrieve food for the ape [61]. Unlike humans, however, non-human primates rarely (if ever) use pointing or other gestures aimed at sharing an attitude about the designated referent (expressive declaratives *sensu* Tomasello [69], e.g. [70,71]).

Most importantly, pointing *for* conspecifics and thus the sharing of information is a rather rare event in non-human primates [72]. There is one report about one incidence of pointing in wild bonobos [73], and some studies with language-trained apes mention the use of pointing gestures in interactions with other apes [65]. However, note that the communicative behaviour of those language-trained apes is largely influenced by their raising history and thus their close proximity to the human culture [66,74]. Therefore, pointing for other conspecifics is extremely rare among wild and captive, non-enculturated apes. The flexible and intentional use of this gesture has been only systematically documented for interactions with humans. Gómez [75] argues that captive non-human primates are restricted by cages and therefore use humans as tools to make them do things for them. Interestingly, an uncaged hand-reared gorilla grabbed the hand of the human and took him to the desired object or target of action and therefore preferred contact gestures instead of pointing [70]. Therefore, it seems unlikely that they simply learn to point by trial and error, but it is suggested that they recruit existing cognitive skills into this referential form of communication [75]. For monkeys, the situation seems to be different, since pointing seems to be ritualized from previously reaching for the food [75].

To summarize, great apes and to some extent also gibbons and monkeys use a variety of manual gestures to communicate with other group members, mostly to request immediate actions of their social partner. Thus, they use their gestures mostly in a dyadic, imperative way to get others to do something for them. Interestingly, Bard [76,77] referred to gestural communication as 'social tool use', which is also reflected in the use of pointing gestures in interactions with humans. Unlike humans, non-human primates do not point for conspecifics and their gestures are often derived from functional actions rather than created as arbitrary ones for communicative purposes ([15], pp. 37–40; [78]), although there are single reports about iconic gestures in gorillas and the use of pantomime in orangutans [79,80]. One possible explanation for those observed differences between different groups of great apes may be based on the different ways in which gestures are acquired. In the next section, we turn our attention to this issue.

(c) *Gesture origins*

Since the focus of this paper is on hand-based gestures, we begin this section with a brief description of how hands are used for the purpose of communication by great apes and monkeys. Hands did not evolve as communicative devices in the first place. In fact, the hands of primates are characterized by an extraordinary degree of primitiveness [81], since the basic, five-fingered appearance is shared not only with other mammals, but even other vertebrates. Still, only in primates does the hand serve a variety of functions including locomotion, manipulation and communication [82]. Moreover, each of these functions is represented by a variety of forms. Thus, locomotion can include things like walking, climbing, jumping or brachiating. Manipulation can include actions such as touching, holding or grasping and more complex forms that combine these basic forms with other more elaborate actions that enable primates to engage in a range of fine-grained activities such as grooming and tool use.

From an evolutionary point of view, Napier ([81], p. 14) noted a '...trend ... to emancipate the hands from weight-bearing to sensitive and delicate multipurpose tools'. However, those different functions are not representative for all primate species but very much depend on the differentiation of the hand in the different taxa. While many monkeys and apes have prehensile hands with nails and in some cases even independently movable or opposable thumbs, other primates such as marmosets and tamarins lack those features. Moreover, the gradual shrinkage of the hands' palmar pads in phylogeny correlates with an increase in prehensility and tactile sensitivity [81]. With the emancipation of forelimbs for manipulatory purposes, the stage is set for the development of hands as communicative devices. Indeed, it is not hard to find potential commonalities between manipulative activity and communicative displays. For instance, monkeys and apes touch, push or pull other's fur during communication. Apes beg for food by placing a cupped hand under the chin of a potential food donor as if to catch food that may fall out. Even in the case of locomotor activity, we can find connections between locomotion and communication.

An intriguing and contentious issue refers to the origin of those communicative displays. One possibility is that they evolved over evolutionary time solely for communicative purposes or that they originally evolved for one function (e.g. locomotion) and were co-opted and reused for a communicative function. Alternatively, communicative gestures may have become ritualized not over evolutionary time but in interactions between individuals and thus over a much shorter time span, an individual's lifetime. Next, we turn our attention to the possible changes involved in gesture origin depending on whether changes take place over evolutionary time (phylogenesis) or an individual's lifetime (ontogenesis).

(d) *Phylogenetic origins*

Animal communication can be very complex and highly ritualized. Perhaps the most famous example is the bee 'language' consisting of different dances to indicate

the position of food resources to other members of the hive [83]. Ritualized communication is not only found in invertebrates. There are many examples from vertebrates, including the complex mating display dance of the stickleback or some lek breeding birds [84,85]. In some cases, the communicative displays are composed of a set of discrete actions that follow a fixed sequence, while in other cases they are constituted by single units. Such signals are displayed by all individuals of the species under a set of predetermined conditions and, critically, they appear even if individuals had no opportunity to observe or interact with other individuals to acquire them.

Whereas some communicative displays seem to have evolved for communicative purposes only, other displays appear to have been 'borrowed' from other contexts and thus from movements that previously had no communicative function via a process called phylogenetic ritualization [86]. For instance, dominance signals such as mounting in monkeys are likely to have evolved from mating behaviour, while some courting displays in birds include elements of foraging behaviour. This *principle of derived activities* [87] refers to actions that originally served a different function but were borrowed and modified to some extent to accomplish a communicative function, in some cases even in a different context from its original one.

If we assume that phylogenetic ritualization is the main mechanism underlying gestural communication in non-human primates, then repertoires of each species should be highly uniform and species-specific gestures should be used even if individuals never had contact with another conspecific. Gestures appear fully formed even when subjects have not had a chance to interact with other individuals. Ground-slapping and chest-beating would be examples of these behaviours [88]. However, that they are phylogenetically ritualized does not mean that they are totally inflexible because, at the very least, they are deployed in the right circumstances and the existence of appropriate substrates/elements determines their appearance. A phylogenetic origin of gestures would mean that all members of a given species should inherit their gestural repertoire, as is the case for vocalizations and, provided with the right conditions, all members of the species would display them. However, it is important to consider that some gestures might be limited to certain developmental stages, resulting in species-typical gestures that are restricted to particular age classes.

(e) Ontogenetic origins

An alternative mechanism for the origin of gestures entails individuals acquiring them during their interactions with conspecifics during their lifetimes rather than inheriting them as postulated above. One such process that involves two individuals mutually shaping each other's behaviour during the course of repeated interactions is called ontogenetic ritualization [29,89]. Initially, individuals use functional behaviours to affect their partner's behaviour. For instance, when they want to embrace their partner, initially they simply pull their partner towards themselves and when they are within reach, they embrace them.

Over repeated interactions, partners begin to anticipate the individual's goal and react before the individual actually has a chance to pull the partner. Next, the individual (anticipating that their partner will react appropriately) does not actually pull but gives an even more abbreviated version of the pull and their partner reacts. Once this stage is reached, we can say that the instrumental action of pulling has become ritualized into a communicative signal.

Ontogenetic ritualization as the main mechanism of gesture acquisition would result in a high degree of variability of individual repertoires and particularly in the occurrence of idiosyncratic gestures, which are exclusive for single individuals only [27]. Idiosyncratic gestures, which were found in all great ape species (for an overview, see [18]), seem to rule out phylogenetic ritualization and thus a genetic determination of an individual's gestural repertoire, since those instances clearly indicate that new gestures can be acquired during an individual's lifetime. Although we still know very little about how such an individually learned new gesture spreads across other group members, there is some evidence that such a transmission takes place, as was shown for the *grooming handclasp* in a captive group of chimpanzees [90].

Variability between groups is evident in the occurrence of group-specific gestures, which are used by the majority of individuals in one group, but are absent in another group. Although group-specific gestures are infrequent, they are reported for chimpanzees [26], gorillas [27,38] and orangutans [28] in captive settings, but also in wild populations, like the *grooming handclasp* of wild chimpanzees [91].

Two basic kinds of gestures have been described in this context: intention movements and attention-getters [92]. Intention movements result from the abbreviations of full-fledged behaviour. For instance, the gesture *arm raise* has been hypothesized to originate from play hitting, initially a functional behaviour that acquires a value as a signal of impending actions. Intention movements typically convey a clear message and are used in a restricted set of social contexts. Moreover, their meaning and origin can be deduced based on use in those contexts. The second kind of gesture is the so-called attention-getter. It is true that the name attention-getter is not very fortunate because unlike what its name suggests, attention-getters are not just designed to capture attention. In fact, their main function may be to trigger others into action, not to call their attention. That they also serve to capture attention may be a by-product. However, there are inconsistent results in terms of whether great apes actually use their gestures to attract the attention of others. In interactions with conspecifics, chimpanzees use either *poke at* or *throw stuff*—both heavily tactile gestures—to attract the attention of the unattending individual [92]. However, this seems to account only for those particular gestures, since further research found that chimpanzees also use auditory gestures more often towards an attentive recipient and tactile gestures were used regardless of the attentional state of the recipient [26,50]. In other words, tactile and auditory gestures are not used particularly often if the recipient is not attending.

However, in interactions with humans, orangutans, gorillas and chimpanzees do use attention-getting behaviours more when they interact with a human who is facing away compared with situations when the human is facing them [53,93–95]. The different findings for interactions between conspecifics on the one hand and interactions with humans on the other hand might also be explained by the constraints of the captive setting. When apes encountered a human with her back turned and they were given a choice between positioning themselves in front of the human or using an auditory gesture to call the human's attention, all great apes species preferred to walk in front of the human to gesture [54]. Thus, similar to the use of pointing gestures, the use of attention-getters might depend very much on the restraints of captivity.

Compared with intention movements, attention-getters appear to be less context-dependent as they appear in multiple contexts for multiple purposes. Additionally, unlike intention movements, it is not easy to envisage a history of ontogenetic ritualization from pre-existing social behaviours as their origin, so that they are possibly also phylogenetic in origin.

There is a second way in which individuals could acquire gestures during ontogeny without requiring ritualization: learning gestures by observation. One possibility is that the individual would copy the gestures that another individual is directing to her (second-person imitation). Another possibility is that the individual could observe two individuals gesturing to each other and acquires those gestures herself without directly interacting with others (third-person imitation). Interestingly, gestures learned by observation walk an opposite path from those that are ontogenetically ritualized. They are acquired fully formed, the individual does not transform an existing behaviour into a streamlined version that becomes the gesture. The individual copies the streamlined version. The resulting outcome would be a high degree of uniformity within the group, paired with substantial differences between groups because each group may have developed their own idiosyncratic gestures and transmitted them across generations. This is clearly the case in humans but it is unclear whether that is also the case in non-human primates.

(f) Phylogenetic versus ontogenetic origins of gestures?

There is currently some debate about what may be the most likely origin of gestures. We have indicated three potential origins for gestures. Historically, observational learning had been proposed as a main mechanism for gesture acquisition. However, there are very little data supporting the idea that apes learn gestures, especially visual gestures, by imitation [26]. Note that the variability in gestural use within groups is as large as between groups. This is not what one would expect based on imitation and cultural transmission since between-group variability should be higher than within-group variability as is the case in humans.

Unlike observational learning, ontogenetic ritualization can explain this pattern of results because the homogeneity within groups would be reduced by the presence of idiosyncratic gestures developed by some individuals but not others. The reason for idiosyncrasy stems from the fact that certain dyadic interactions between individuals are unique, for instance, mothers and infants may follow different caregiving routines. In fact, according to Tomasello *et al.* [29], the presence of idiosyncratic gestures is a key indicator of ontogenetic ritualization and evidence against a phylogenetic origin of gestures.

Genty *et al.* [37] have recently challenged this idea and proposed that ape gestures are not ontogenetically ritualized but appear fully formed in individuals. This phylogenetic origin of gestures leaves no room for modification over time. They argue that the differences between groups and the idiosyncrasy that has been described are a consequence of the sampling methods that have been used. In particular, not enough hours have been observed to be able to obtain the whole repertoire of gestures. Thus, idiosyncrasy results from a low sampling effort as opposed to individualized experiences with other conspecifics. An increase in the sampling effort showed that all individuals used the same gestures and virtually eliminated idiosyncratic gestures from the sample. This result casts some doubt not only on ontogenetic ritualization but also on observational learning as acquisition mechanisms because there were no differences between groups.

However, one limitation of this and many other studies is that they are not longitudinal and therefore they cannot detect change either within an individual's lifespan or across generations. So, although all individuals use the same gestures, this does not prove that gestures have not undergone an ontogenetic ritualization process. What is needed are long-term studies actually investigating whether the gestures of great apes (and monkeys) change over time. Additionally, studies that have investigated gestural acquisition of apes in contact with humans have described the ritualization process [96]. One could argue that apes in contact with humans would be different, but this is hard to reconcile with the idea that human-reared apes were requesting the same things (e.g. go to another location) as the ones living with their biological mothers. Nevertheless, it is true that the case for ontogenetic ritualization may have been overstated because as Genty *et al.* [37] point out, it is difficult to envisage a history of ontogenetic ritualization for some gestures such as chest beating or ground slapping, although it is also true that ontogenetic ritualization may still be a viable alternative for other gestures such as gentle touch or arm raise.

After discussing the origins of gestures in non-human primates and their close link to actions, we will now briefly refer to some of the current theories on language evolution and the role gestures might have played, before we address the question of laterality in gesture use in non-human primates.

3. LANGUAGE ORIGINS (OUT OF ACTIONS)
(a) Gestural origin of human language
The origin of human language is a fiercely debated question, with some scholars favouring a vocal origin (e.g. [97]) and others emphasizing gestures as

precursors to human language (e.g. [43]). To our knowledge, there is no coherent theory currently available that has attempted to reconcile the two opposing sets of theories, which usually see themselves as mutually exclusive [46]. Gestural theories usually refer to the very flexible use of gestures across different contexts and the fact that new gestures can be learned and incorporated into a species repertoire [98]. The discovery of a mirror-neuron system for grasping in monkeys [99] has nourished a variety of evolutionary scenarios focusing on the role of gestures in this process, since mirror neurons represent the link between manual, practical actions and communication.

(b) *From action to language*
Mirror neurons allow macaques to establish a link between performing an action and being able to recognize it [99]. Interestingly, these neurons are located in a brain area that is homologous to Broca's area in the human brain responsible for processing language. Mirror neurons therefore might have played an important role in the evolution of human language, since they were already present in our ancestors representing the neural prerequisite for the development of inter-individual communication and finally of speech [100]. Next, we will give a brief overview of theories suggesting a gestural origin, and second, we will turn to the lateralization of hand use while gesturing and the evidence currently available for non-human primates.

(c) *Gesture and laterality*
Gestural theories of language evolution often refer to the link between lateralization of hand use and language [43]. In humans, the motor systems controlling both manual and oral movements are usually lateralized to the left hemisphere [101]. Therefore, the majority of the human population is right-handed, with the left hemisphere controlling movements of the right hand. Furthermore, both language production and comprehension are located in distinct areas of the left brain hemisphere [102]. The close link between language and manual actions becomes evident in studies showing that while speaking, humans gesture significantly more with their right hand compared with their left hand [103]. This suggests that the functional asymmetry is not specific for one modality only, and that the production of speech apparently also activates motor areas in the left hemisphere, resulting in an increased use of the right hand [104].

Comparable evidence has been found for non-human primates since they show a preference for using their right hand for different manual actions including gestures while vocalizing [105,106]. These findings suggest that the lateralization of manual and oral movements represents a trait shared by both humans and other primates.

However, results for the preference for one hand and particularly the right hand are not completely consistent [107,108]. Although a hand preference is found in many monkey and ape species for different manual actions such as carrying, tool use and locomotion [105,109–111], hand preference is often task-specific and often only evident on an individual, but not

species level [112,113]. So far, there is little evidence that gestures used to communicate with conspecifics—and thus not in interactions with humans when begging for food—are mainly produced with the right hand. To our knowledge, there is only one study showing that baboons use their right hand while they gesture, but not when they perform non-communicative actions [22]. In interactions with humans, however, there is clear evidence that chimpanzees use their right hand preferentially while gesturing [114–116], and they also used their right hand significantly more while producing gestures compared with other manual actions. Hopkins *et al.* [117] therefore concluded that the left-hemisphere specialization for language may have evolved initially from asymmetries in manual gestures in the common ancestor of chimpanzees and humans, rather than from hand use associated with other, non-communicative motor actions such as tool use and bimanual actions. The laterality of hand use in chimpanzees is also reflected in neuroanatomical asymmetries, since chimpanzees that preferably gesture with their right hand also have larger inferior frontal gyri in the left hemisphere than those apes that do not show consistent hand use while gesturing [118].

To summarize, those studies indicate that manual gestures of at least chimpanzees are lateralized, and this functional asymmetry is also associated with asymmetries in the corresponding neural substrate. Hopkins *et al.* [117] therefore suggest that the dominance of the left hemisphere for language has evolved from a gestural communication system already lateralized in the left hemisphere in our common ancestor 5–7 Ma.

4. CONCLUSION
In our contribution, we wanted to highlight that the hands of non-human primates, and particularly those of the great apes, are suitable tools to perform a variety of gestures of different modalities. They are used to achieve a range of different social goals and display a high degree of flexibility as indicated by the possibility of acquiring new gestures often outside of what would be the species-specific repertoire. Still, gestures of non-human primates are different from human gestures in many aspects, since they are mostly used in a dyadic and imperative way, and they also lack the high level of abstraction typical for human gestures. Thus, gestures of non-human primates may emerge from actions via three potential pathways.

The high degree of variability between individual repertoires, the occurrence of idiosyncratic gestures and thus the creation of new gestures support the idea that ontogenetic ritualization may be involved in the origin of some gestures. However, other gestures appear more or less fully formed even in the absence of conspecifics, thus indicating a strong genetic predisposition to develop certain gestures. Finally, some form of social learning might also be implicated either in the form of facilitating the appearance of some gestures or perhaps even the acquisition of novel gestures, although this still needs to be supported by empirical evidence.

After elaborating on the close link between manual actions and gestural communication, we turned to the question of whether gestures of non-human primates are lateralized as many human manual actions including certain gesture types are. There is some evidence for right-handedness at least in captive chimpanzees, and, interestingly, they use their right hand even more while vocalizing, thus suggesting a close link between the manual and oral movements. This fact is often used to support a gestural origin of human language, since the functional asymmetry of hand use while gesturing is also present in the neural substrate of chimpanzees, suggesting some continuity in our phylogenetic history. However, one must keep in mind that the evidence of laterality in chimpanzees and other non-human primate species at the population level is quite mixed. This means that it may be too early to generalize a right-hand preference for gesture use in our closest relatives.

There is much to be done in the future to trace the origins of gestures. Longitudinal studies are especially important as they can throw light on how gestures actually emerge in both monkeys and apes. Some research effort devoted to non-great ape species would be particularly welcome. Otherwise the field runs the risk of underestimating what aspects of gestural communication that are common to human and non-human apes are already present in monkeys. Finally, there is much work to be done in terms of unifying concepts and criteria across the various disciplines that conduct research on gestural communication.

REFERENCES

1 Gómez, J. C. 1990 The emergence of intentional communication as a problem-solving strategy in gorilla. In *'Language' and intelligence in monkeys and apes* (eds S. T. Parker & K. R. Gibson), pp. 333–355. Cambridge, UK: Cambridge University Press.

2 Simpson, J. A. & Weiner, E. S. C. (eds) 1998 The Oxford English Dictionary, 2nd edn. Oxford, UK: Oxford University Press.

3 Kendon, A. 2004 *Gesture: visible action.* Cambridge, UK: Cambridge University Press.

4 Liszkowski, U., Carpenter, M., Henning, A., Striano, T. & Tomasello, M. 2004 Twelve-months-olds point to share attention and interest. *Dev. Sci.* **7**, 297–307. (doi:10.1111/j.1467-7687.2004.00349.x)

5 Butterworth, G. & Grover, L. 1988 The origins of referential communication in human infancy. In *Thought without language* (ed. L. Weiskrantz), pp. 5–24. Oxford, UK: Clarendon Press.

6 Franco, F. & Butterworth, G. 1996 Pointing and social awareness: declaring and requesting in the second year. *J. Child Lang.* **23**, 307–336. (doi:10.1017/S0305000900008813)

7 Carpenter, M., Nagell, K. & Tomasello, M. 1998 Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monogr. Soc. Res. Child Dev.* **63**, 176. (doi:10.2307/1166214)

8 Volterra, V., Caselli, M. C., Caprici, O. & Pizzuto, E. 2005 Gesture and the emergence and development

of language. In *Elizabeth Bates: a festschrift* (eds M. Tomasello & D. Slobin). Mahwah, NJ: Lawrence Erlbaum Associates.

9 Capirci, O., Montanari, S. & Volterra, V. 1998 Gestures, signs, and words in early language development. *New Dir. Child Dev.* **79**, 45–60.

10 Alibali, M. W., Kita, S. & Young, A. J. 2000 Gesture and the process of speech production: we think, therefore we gesture. *Lang. Cogn. Proc.* **15**, 593–613. (doi:10.1080/016909600750040571)

11 Goldin-Meadow, S. 2002 Constructing communication by hand. *Cogn. Dev.* **17**, 1385–1405. (doi:10.1016/S0885-2014(02)00122-3)

12 McNeill, D. 2000 *Language and gesture.* Cambridge, UK: Cambridge University Press.

13 Goldin-Meadow, S. 2003 *The resilience of language: what gesture creation in deaf children can tell us about how all children learn language.* New York, NY: Psychology Press.

14 Senghas, A., Kita, S. & Özyürek, A. 2004 Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* **305**, 1779–1782. (doi:10.1126/science.1100199)

15 McNeill, D. 1992 *Hand and mind: what gestures reveal about thought.* Chicago, IL: University of Chicago Press.

16 Bates, E. 1979 *The emergence of symbols: cognition and communication in infancy.* New York, NY: Academic Press.

17 Bates, E., Camaioni, L. & Volterra, V. 1975 The acquisition of performatives prior to speech. *Merrill-Palmer Quart.* **21**, 205–226.

18 Call, J. & Tomasello, M. (eds) 2007 *The gestural communication of apes and monkeys.* Mahwah, NJ: Lawrence Erlbaum Associates.

19 Leavens, D. A., Russell, J. L. & Hopkins, W. D. 2005 Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Dev.* **76**, 291–306. (doi:10.1111/j.1467-8624.2005.00845.x)

20 Maestripieri, D. 1999 Primate social organization, gestural repertoire size, and communication dynamics. In *The origins of language: what nonhuman primates can tell us* (ed. B. King), pp. 55–77. Santa Fe, NM: School of American Research Press.

21 Paukner, A., Anderson, J. R., Fogassi, L. & Ferrari, P. F. 2006 Do facial gestures, visibility or speed of movement influence gaze following responses in pigtail macaques? *Primates* **48**, 241–244. (doi:10.1007/s10329-006-0024-z)

22 Meguerditchian, A. & Vauclair, J. 2006 Baboons communicate with their right hand. *Behav. Brain Res.* **171**, 170–174. (doi:10.1016/j.bbr.2006.03.018)

23 Hesler, N. & Fischer, J. 2007 Gestural communication in Barbary macaques (*Macaca sylvanus*): an overview. In *The gestural communication of apes and monkeys* (eds J. Call & M. Tomasello), pp. 159–196. Mahwah, NJ: Lawrence Erlbaum Associates.

24 Laidre, M. E. 2008 Do captive mandrills invent new gestures? *Anim. Cogn.* **11**, 179–187. (doi:10.1007/s10071-007-0121-4)

25 Liebal, K., Pika, S. & Tomasello, M. 2004 Social communication in siamangs (*Symphalangus syndactylus*): use of gestures and facial expressions. *Primates* **45**, 41–57. (doi:10.1007/s10329-003-0063-7)

26 Tomasello, M., Call, J., Warren, J., Frost, G. T., Carpenter, M. & Nagell, K. 1997 The ontogeny of chimpanzee gestural signals: a comparison across groups and generations. *Evol. Commun.* **1**, 223–259. (doi:10.1075/eoc.1.2.04tom)

27 Pika, S., Liebal, K. & Tomasello, M. 2003 Gestural communication in young gorillas (*Gorilla gorilla*):

gestural repertoire, learning, and use. *Am. J. Primatol.* **60**, 95–111. (doi:10.1002/ajp.10097)

28 Liebal, K., Pika, S. & Tomasello, M. 2006 Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* **6**, 1–38. (doi:10.1075/gest.6.1.02lie)

29 Tomasello, M., George, B., Kruger, A., Farrar, J. & Evans, E. 1985 The development of gestural communication in young chimpanzees. *J. Hum. Evol.* **14**, 175–186. (doi:10.1016/S0047-2484(85)80005-1)

30 MacKinnon 1974 The behaviour and ecology of wild orang-utans (*Pongo pygmaeus*). *Anim. Behav.* **22**, 3–74. (doi:10.1016/S0003-3472(74)80054-0)

31 Schaller, G. B. 1963 *The mountain gorilla: ecology and behavior.* Chicago, IL: Chicago University Press.

32 Goodall, J. 1986 *The chimpanzees of Gombe: patterns of behavior.* Cambridge, MA: Harvard University Press.

33 Kuroda, S. J. 1980 Social behavior of the pygmy chimpanzees. *Primates* **21**, 181–197. (doi:10.1007/BF02374032)

34 Baldwin, L. A. & Teleki, G. 1976 Patterns of gibbon behavior on Hall's Island, Bermuda: a preliminary ethogram for *Hylobates lar*. In *Gibbon and siamang* (ed. D. Rumbaugh), pp. 21–105. Basel, Switzerland: Karger.

35 Chivers, D. 1976 Communication within and between family groups of siamang (*Symphalangus syndactylus*). *Behaviour* **57**, 116–135. (doi:10.1163/156853976X00136)

36 Kummer, H. 1968 *Social organization of hamadryas baboons. A field study.* Basel, Switzerland: Karger.

37 Genty, E., Breuer, T., Hobaiter, C. & Byrne, R. 2009 Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim. Cogn.* **12**, 527–546. (doi:10.1007/s10071-009-0213-4)

38 Tanner, J. & Byrne, R. 1999 The development of spontaneous gestural communication in a group of zoo-living lowland gorillas. In *The mentalities of gorillas and orangutans* (eds T. Parker, S. Milks & R. Mitchell), pp. 211–239. Cambridge, UK: Cambridge University Press.

39 Maestripieri, D. 1996 Gestural communication and its cognitive implications in pigtail macaques (*Macaca nemestrina*). *Behaviour* **133**, 997–1022. (doi:10.1163/156853996X00576)

40 Maestripieri, D. 1996 Social communication among captive stump-tailed macaques (*Macaca arctoides*). *Int. J. Primatol.* **17**, 785–802. (doi:10.1007/BF02735264)

41 Maestripieri, D. 1997 Gestural communication in macaques: usage and meaning of nonvocal signals. *Evol. Commun.* **1**, 193–222. (doi:10.1075/eoc.1.2.03mae)

42 Arbib, M. A., Liebal, K. & Pika, S. 2008 Primate vocalization, gesture, and the evolution of human language. *Curr. Anthropol.* **49**, 1053–1076. (doi:10.1086/593015)

43 Corballis, M. C. 2003 From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* **26**, 199–260. (doi:10.1017/S0140525X03000050)

44 Hewes, G. W. 1973 Primate communication and the gestural origin of language. *Curr. Anthropol.* **12**, 5–24. (doi:10.1086/201401)

45 Pollick, A. S. & de Waal, F. 2007 Ape gestures and language evolution. *Proc. Natl Acad. Sci. USA* **104**, 8184–8189. (doi:10.1073/pnas.0702624104)

46 Slocombe, K., Waller, B. & Liebal, K. 2011 The language void: the need for multimodality. *Anim. Behav.* **81**, 919–924. (doi:10.1016/j.anbehav.2011.02.002)

47 Cartmill, E. A. & Byrne, R. W. 2010 Semantics of primate gestures: intentional meanings of orangutan gestures. *Anim. Cogn.* **13**, 793–804. (doi:10.1007/s10071-010-0328-7)

48 Pika, S. 2008 What is the nature of the gestural communication of great apes? In *The shared mind: perspectives on intersubjectivity* (eds J. Zlatev, T. P. Racine, C. Sinha & E. Itkonen), pp. 165–186. Amsterdam, The Netherlands: John Benjamins Publishing Company.

49 Genty, E. & Byrne, R. 2010 Why do gorillas make sequences of gestures? *Anim. Cogn.* **13**, 287–301. (doi:10.1007/s10071-009-0266-4)

50 Liebal, K., Call, J. & Tomasello, M. 2004 Use of gesture sequences in chimpanzees (*Pan troglodytes*). *Am. J. Primatol.* **64**, 377–396. (doi:10.1002/ajp.20087)

51 Tanner, J. 2004 Gestural phrases and gestural exchanges by a pair of zoo-living lowland gorillas. *Gesture* **4**, 1–24. (doi:10.1075/gest.4.1.02tan)

52 Rosati, A. G. & Hare, B. 2009 Looking past the model species: diversity in gaze-following skills across primates. *Curr. Opin. Neurobiol.* **19**, 45–51. (doi:10.1016/J.Conb.2009.03.002)

53 Hostetter, A. B., Cantero, M. & Hopkins, W. D. 2001 Differential use of vocal and gestural communication by chimpanzees (*Pan troglodytes*) in response to the attentional status of a human (*Homo sapiens*). *J. Comp. Psychol.* **115**, 337–343. (doi:10.1037//0735-7036.115.4.337)

54 Liebal, K., Pika, S., Call, J. & Tomasello, M. 2004 To move or not to move: how great apes adjust to the attentional state of others. *Interact. Stud.* **5**, 199–219. (doi:10.1075/is.5.2.03lie)

55 Tempelmann, S. & Liebal, K. In press. Spontaneous use of gesture sequences in orangutans. A case for strategy? *Gesture*.

56 Cartmill, E. A. & Byrne, R. W. 2007 Orangutans modify their gestural signaling according to their audience's comprehension. *Curr. Biol.* **17**, 1345–1348. (doi:10.1016/j.cub.2007.06.069)

57 Leavens, D. A., Hopkins, W. D. & Bard, K. A. 1996 Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **110**, 346–353. (doi:10.1037/0735-7036.110.4.346)

58 Mitchell, R. & Anderson, J. 1997 Pointing, withholding information, and deception in capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* **111**, 351. (doi:10.1037/0735-7036.111.4.351)

59 Anderson, J., Kuwahata, H. & Fujita, K. 2007 Gaze alternation during 'pointing' by squirrel monkeys (*Saimiri sciureus*)? *Anim. Cogn.* **10**, 267–271. (doi:10.1007/s10071-006-0065-0)

60 Call, J. & Tomasello, M. 1994 The production and comprehension of referential pointing by orangutans (*Pongo pygmaeus*). *J. Comp. Psychol.* **108**, 307–317. (doi:10.1037/0735-7036.108.4.307)

61 Zimmermann, F., Zemke, F., Call, J. & Gómez, J. C. 2009 Orangutans (*Pongo pygmaeus*) and bonobos (*Pan paniscus*) point to inform a human about the location of a tool. *Anim. Cogn.* **12**, 347–358. (doi:10.1007/s10071-008-0194-8)

62 Leavens, D. A. & Hopkins, W. D. 1998 Intentional communication by chimpanzees (*Pan troglodytes*): a cross-sectional study of the use of referential gestures. *Dev. Psychol.* **34**, 813–822. (doi:10.1037/0012-1649.34.5.813)

63 Leavens, D. A., Hopkins, W. D. & Thomas, R. K. 2004 Referential communication by chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **118**, 48–57. (doi:10.1037/0735-7036.118.1.48)

64 Pedersen, J. & Fields, W. M. 2009 Aspects of repetition in bonobo–human conversation: creating cohesion in a conversation between species. *Integr. Psychol. Behav.* **43**, 22–41. (doi:10.1007/s12124-008-9067-6)

65 Savage-Rumbaugh, E. S. 1986 *Ape language: from conditioned response to symbol.* New York, NY: Columbia University Press.

66 Leavens, D. A., Hopkins, W. D. & Bard, K. A. 2005 Understanding the point of chimpanzee pointing.

Curr. Dir. Psychol. Sci. **14**, 185–189. (doi:10.1111/j.0963-7214.2005.00361.x)

67 Liszkowski, U., Carpenter, M., Striano, T. & Tomasello, M. 2006 Twelve and 18-months-olds point to provide information for others. *J. Cogn. Dev.* **7**, 297–307.

68 Call, J. 2011 How artificial communication affects the communication and cognition of the great apes. *Mind Lang.* **26**, 1–20. (doi:10.1111/j.1468-0017.2010.01408.x)

69 Tomasello, M. 2008 *Origins of human communication.* Cambridge, MA: MIT Press.

70 Gómez, J. C., Sarria, E. & Tamarit, J. 1993 The comparative study of early communication and theories of mind: ontogeny, phylogeny, and pathology. In *Understanding other minds: perspectives from autism* (eds S. Baron-Cohen & H. Tager-Flusberg), pp. 397–426. Oxford, UK: Oxford University Press.

71 Tomonaga, M., Tanaka, M., Matsuzawa, T., Myowa-Yamakoshi, M., Kosugi, D., Mizuno, Y., Okamoto, S., Yamaguchi, M. K. & Bard, K. 2004 Development of social cognition in infant chimpanzees (*Pan troglodytes*): face recognition, smiling, gaze, and the lack of triadic interactions. *Jpn Psychol. Res.* **46**, 227–235. (doi:10.1111/j.1468-5584.2004.00254.x)

72 Gómez, J. C. 2007 Pointing behaviors in apes and human infants: a balanced interpretation. *Child Dev.* **78**, 729–734. (doi:10.1111/j.1467-8624.2007.01027.x)

73 Vea, J. J. & Sabater-Pi, J. 1998 Spontaneous pointing behaviour in the wild pygmy chimpanzee (*Pan paniscus*). *Folia Primatol. (Basel)* **69**, 289–290. (doi:10.1159/000021640)

74 Cartmill, E. A. & Maestripieri, D. 2012 Socio-cognitive specializations of nonhuman primates: evidence from gestural communication. In *The Oxford hand book of comparative evolutionary psychology* (eds J. Vonk & T. Shackelford). Oxford, UK: Oxford University Press.

75 Blaschke, M. & Ettlinger, G. 1987 Pointing as an act of social communication by monkeys. *Anim. Behav.* **35**, 1520–1523. (doi:10.1016/S0003-3472(87)80023-4)

76 Bard, K. 1990 'Social tool use' by free-ranging orangutans: a Piagetian and developmental perspective on the manipulation of an animate object. In *'Language' and intelligence in monkeys and apes: comparative developmental perspectives* (eds S. T. Parker & K. R. Gibson), pp. 356–378. New York, NY: Cambridge University Press.

77 Bard, K. 1992 Intentional behaviour and intentional communication in young free-ranging orangutans. *Child Dev.* **63**, 1186–1197. (doi:10.1111/1467-8624.ep9301210043)

78 Kendon, A. 1988 How gestures can become like words. In *Cross-cultural perspectives in nonverbal communication* (ed. F. Poyatos), pp. 131–141. Toronto, ON: Hogrefe & Huber.

79 Russon, A. & Andrews, K. 2010 Orangutan pantomime: elaborating the message. *Biol. Lett.* (doi:10.1098/rsbl.2010.0564)

80 Tanner, J. & Byrne, R. 1996 Representation of action through iconic gesture in a captive lowland gorilla. *Curr. Anthropol.* **37**, 162–173. (doi:10.1086/204484)

81 Napier, J. R. 1976 *The Human hand.* Burlington, NC: Carolina Biological Supply Company.

82 Schultz, A. H. 1968 Form und Funktion der Primatenhände. In *Handgebrauch und Verständigung bei Affen und Frühmenschen* (ed. B. Rentsch), pp. 9–30. Bern, Switzerland: Hans Huber.

83 Frisch, K. V. 1973 Honeybees: do they use direction and distance information provided by their dancers? In *Perception: an adaptive process* (ed. T. L. Bennett), pp. 84–91. New York, NY: MSS Information Corp.

84 Kirkpatrick, M. & Ryan, M. 1991 The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33–38. (doi:10.1038/350033a0)

85 Tinbergen, N. 1954 The origin and evolution of courtship and threat display. In *Evolution as a process* (eds A. C. Hardy, J. S. Huxley & E. B. Ford), pp. 233–250. London, UK: Allen and Unwin.

86 Krebs, J. R. & Dawkins, R. 1984 Animal signals: mind-reading and manipulation. In *Behavioral ecology: an evolutionary approach* (eds J. R. Krebs & N. B. Davies), pp. 380–402. Oxford, UK: Blackwell Scientific Publications.

87 Tinbergen, N. 1952 'Derived' activities; their causation, biological significance, origin, and emancipation during evolution. *Q. Rev. Biol.* **27**, 1–32. (doi:10.1086/398642)

88 Redshaw, M. & Locke, K. 1976 The development of play and social behaviour in two lowland gorilla infants. *Dodo J. Jersey Wildlife Preserv. Trust* **13**, 71–86.

89 Plooij, F. X. 1978 Some basic traits of language in wild chimpanzees? In *Action, gesture and symbol: the emergence of language* (ed. A. Lock), pp. 111–131. London, UK: Academic Press.

90 Bonnie, K. & de Waal, F. 2006 Affiliation promotes the transmission of a social custom: handclasp grooming among captive chimpanzees. *Primates* **47**, 27–34. (doi:10.1007/s10329-005-0141-0)

91 McGrew, W. C. & Tutin, C. E. G. 1978 Evidence for a social custom in wild chimpanzees? *Man* **13**, 234–251. (doi:10.2307/2800247)

92 Tomasello, M., Gust, D. & Frost, G. 1989 A longitudinal investigation of gestural communication in young chimpanzees. *Primates* **30**, 35–50. (doi:10.1007/BF02381209)

93 Poss, S., Kuhar, C., Stoinski, T. S. & Hopkins, W. D. 2006 Differential use of attentional and visual communicative signaling by orangutans (*Pongo pygmaeus*) and gorillas (*Gorilla gorilla*) in response to the attentional status of a human. *Am. J. Primatol.* **68**, 978–992. (doi:10.1002/ajp.20304)

94 Hopkins, W. D., Taglialatela, J. P. & Leavens, D. A. 2007 Chimpanzees differentially produce novel vocalizations to capture the attention of a human. *Anim. Behav.* **73**, 281–286. (doi:10.1016/j.anbehav.2006.08.004)

95 Leavens, D. A., Hostetter, A. B., Wesley, M. J. & Hopkins, W. D. 2004 Tactical use of unimodal and bimodal communication by chimpanzees *Pan troglodytes. Anim. Behav.* **67**, 467–476. (doi:10.1016/j.anbehav.2003.04.007)

96 Gómez, J. C. 1990 The emergence of intentional communication as a problem-solving strategy in the gorilla. In *'Language' and intelligence in monkeys and apes: comparative developmental perspectives* (eds S. T. Parker & K. R. Gibson), pp. 333–355. Cambridge, UK: Cambridge University Press.

97 Zuberbühler, K. 2005 The phylogenetic roots of language: evidence from primate communication and cognition. *Curr. Dir. Psychol. Sci.* **14**, 126–130. (doi:10.1111/j.0963-7214.2005.00357.x)

98 Tomasello, M. 2007 Ape gestures and the origins of language. In *Gestural communication in apes and monkeys* (eds J. Call & M. Tomasello), pp. 221–239. Hillsdale, NJ: Lawrence Earlbaum Associates.

99 Rizzolatti, G. & Arbib, M. A. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194. (doi:10.1016/S0166-2236(98)01260-0)

100 Arbib, M. A. 2005 From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**, 105–124. (doi:10.1017/S0140525X05000038)

101 Kimura, D. 1993 *Neuromotor mechanisms in human communication.* New York, NY: Oxford University Press.

102 Corballis, M. C. 1992 *The lopsided brain: evolution of the generative mind.* New York, NY: Oxford University Press.

103 Kimura, D. 1973 Manual activity during speaking—I. Right-handers. *Neuropsychologia* **11**, 45–50. (doi:10. 1016/0028-3932(73)90063-8)

104 Kinsbourne, M. & Hicks, R. E. 1978 Functional cerebral space: a model for overflow, transfer, and interference effects in human performance. In *Attention and performance* (ed. J. Requin), pp. 54–68. Hillsdale, NJ: Erlbaum.

105 Stafford, D. K., Milliken, G. W. & Ward, J. P. 1990 Lateral bias in feeding and brachiation in *Hylobates*. *Primates* **31**, 407–414. (doi:10.1007/BF02381111)

106 Hopkins, W. D. & Cantero, M. 2003 From hand to mouth in the evolution of language: the influence of vocal behavior on lateralized hand use in manual gestures by chimpanzees (*Pan troglodytes*). *Dev. Sci.* **6**, 55–61. (doi:10.1111/1467-7687.00254)

107 McGrew, W. C. & Marchandt, L. F. 1992 Chimpanzees, tools, and termites: hand preference or handedness. *Curr. Anthropol.* **33**, 114–119. (doi:10.1086/204041)

108 Cashmore, L., Uomini, N. & Chapelain, A. 2008 The evolution of handedness in humans and great apes: a review and current issues. *J. Anthropol. Sci.* **86**, 7–35.

109 Westergaard, G. & Suomi, S. 1993 Hand preference in capuchin monkeys varies with age. *Primates* **34**, 295–299. (doi:10.1007/BF02382624)

110 Hopkins, W. D., Bennett, A. J. & Bales, S. L. 1993 Behavioural laterality in captive bonobos (*Pan paniscus*). *J. Comp. Psychol.* **107**, 403–410. (doi:10.1037/0735-7036.107.4.403)

111 Hopkins, W. D., Bard, K. A., Jones, A. & Bales, S. L. 1993 Chimpanzee hand preference in throwing and infant cradling: implications for the origin of human handedness. *Curr. Anthropol.* **34**, 786–790. (doi:10. 1086/204224)

112 Sugiyama, Y., Fushimi, T., Sakura, O. & Matsuzawa, T. 1993 Hand preference and tool use in wild chimpanzees. *Primates* **34**, 151–159. (doi:10.1007/BF02 381386)

113 Anderson, J., Degiorgio, C., Lamarque, C. & Fagot, J. 1996 A multi-task assessment of hand lateralization in capuchin monkeys (*Cebus apella*). *Primates* **37**, 97–103. (doi:10.1007/BF02382926)

114 Hopkins, W. D. & Cantalupo, C. 2005 Individual and setting differences in the hand preferences of chimpanzees *Pan troglodytes*: a critical analysis and some alternative explanations. *Laterality* **10**, 65–80.

115 Hopkins, W. D. 1996 Chimpanzee handedness: causes and consequences. *Int. J. Primatol.* **31**, 3665–3665.

116 Hopkins, W. D. & Leavens, D. A. 1998 Hand use and gestural communication in chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **112**, 95–99. (doi:10.1037/ 0735-7036.112.1.95)

117 Hopkins, W. D., Russell, J., Freeman, H. D., Buehler, N., Reynolds, E. & Schapiro, S. J. 2005 The distribution and development of handedness for manual gestures in captive chimpanzees (*Pan troglodytes*). *Psychol. Sci.* **16**, 487–493(487). (doi:10.1111/j.0956-7976.2005. 01561.x)

118 Taglialatela, J. P., Cantalupo, C. & Hopkins, W. D. 2006 Gesture handedness predicts asymmetry in the chimpanzee inferior frontal gyrus. *Neuroreport* **17**, 923–927. (doi:10.1097/01.wnr.0000221835.26093.5e)

Review

A word in the hand: action, gesture and mental representation in humans and non-human primates

Erica A. Cartmill*, Sian Beilock and Susan Goldin-Meadow

Department of Psychology, University of Chicago, 5848 South University Avenue, Chicago, IL 60637, USA

The movements we make with our hands both reflect our mental processes and help to shape them. Our actions and gestures can affect our mental representations of actions and objects. In this paper, we explore the relationship between action, gesture and thought in both humans and non-human primates and discuss its role in the evolution of language. Human gesture (specifically representational gesture) may provide a unique link between action and mental representation. It is kinaesthetically close to action and is, at the same time, symbolic. Non-human primates use gesture frequently to communicate, and do so flexibly. However, their gestures mainly resemble incomplete actions and lack the representational elements that characterize much of human gesture. Differences in the mirror neuron system provide a potential explanation for non-human primates' lack of representational gestures; the monkey mirror system does not respond to representational gestures, while the human system does. In humans, gesture grounds mental representation in action, but there is no evidence for this link in other primates. We argue that gesture played an important role in the transition to symbolic thought and language in human evolution, following a cognitive leap that allowed gesture to incorporate representational elements.

Keywords: gesture; mental representation; evolution of language; embodied cognition; primates; mirror neurons

1. INTRODUCTION

A growing body of evidence suggests that movements of the body not only reflect processes of the mind but can also influence them. We focus here on one particular type of movement—representational gesture. These gestures have the potential to provide a link between action and thought because gesture offers a vehicle not only for representing information about action, but also for representing that information outside of the context of real-world acts. Representational gestures are hand movements that often resemble the actual movements involved in acting on objects (e.g. rotating the hand in the air as though twisting a jar lid). However, gestures *represent* rather than *replicate* actions. Unlike actions, gestures do not bring about physical change in the environment (the twisting motion does not actually open the lid). They can, however, change how we (and others) think and speak, and may have played a central role in developing the human ability to think and speak.

In this paper, we review and discuss the relationship between action, gesture and mental representation in humans, and assess the comparative evidence for a link between action, gesture and thought in

non-human primates. We begin by reviewing studies of action's influence on thought, focusing on evidence that experience doing actions affects the mental representations of those actions. We then turn to a discussion of human gesture. After describing how and when gesture is used, we address gesture's relationship to action. Representational gestures can depict action in a number of different ways—for example, the hand can represent the hand of an agent performing an action on an object, the object itself or the trajectory of the motion. We end our discussion of human gesture by exploring gesture's influence on mental representation, and conclude that gesture can have a stronger influence on thought than action itself. In fact, it has a unique ability to act as a bridge between thought and action because it is both kinaesthetically close to action and yet also symbolic.

In the final section, we turn to action representation and gesture in non-human primates (specifically, monkeys and apes). We review evidence that non-human primates are adept at understanding and performing actions, but suggest that they are not able to represent actions symbolically in gesture. We focus on the naturalistic (i.e. not human-taught) gestures of great apes and compare them with human gestures. Although there are many differences between human and ape gestures (synchronization with vocalization, systematic patterning, social motivation to share information), it is the ability to represent action through gesture that seems to be

unique to humans and key to the differences in non-vocal communication. Non-human primates can use gesture in complex ways, but their gestures are often abbreviated versions of actions, and are not representational or 'symbolic' in the way that many human gestures are. Moreover, unlike the human mirror neuron system, the monkey mirror system does not respond to manual representational gestures, suggesting that the mirror system may play an important role in distinguishing the way action and gesture are processed in humans versus non-human primates. We conclude that gesture does not serve as a bridge between action and cognition in non-human primates and that building this bridge may have been an important step in human evolution.

2. ACTION AND THOUGHT

(a) *The relation between action and thought*

The mind and the body have historically been studied as separate entities, leading to the view that cognition and action are independent domains (see [1,2] for discussion). Recent theories of human cognition suggest that the mind is not an isolated system but rather is integrated into the body's sensorimotor systems, and that our representations of objects and events are linked to our experiences of acting on the world (e.g. [3–7]). This embodied approach to cognition places a heavy emphasis on the idea that our mental representations of objects, events and many types of information we encounter arise from (and are linked to) our physical experiences interacting with the world. For example, when asked to make preference judgements between non-sensical letter sequences, skilled typists preferred pairs of letters that are typed with different fingers on different hands (letters that would be physically easier to type if one were to type them) than letters typed with the same finger. Novice typists with little previous typing experience had no such preference [8]. Importantly, neither group could explain the differences between the letter pairs, suggesting that the skilled typists' preference was unconsciously based on their previous motor experience of how easy or hard it was to type the presented letters. All else being equal, we generally prefer what is easiest to act on, perceive, read, etc. Our prior sensorimotor experiences are so tightly linked to our mental representations that they can influence our thinking about objects or events even in scenarios far removed from relevant actions (e.g. [8]).

Despite growing evidence that action influences thought, the process through which action interacts with representation is not well understood. Some propose that neural representations of objects and events are built upon neural activations that arose during past experience interacting with objects and events in the world (e.g. [3]). Under this view of embodied cognition, mental representations of objects and events reflect, and to some extent rely on, traces of neural activation (or 'perceptual symbols') caused by past real-world interactions (for a review, see [6]). Others propose that the physical limitations of brain size require that areas primarily responsible for one type of ability be reused for a range of purposes and that these overlaps are primarily responsible for neural

co-activation during physical and mental tasks [9]. Under this view, concepts of particular objects need not be grounded in the actions a person has performed on that object, but rather in the exaptation of one area of the brain for use in another area. For example, finger sensitivity and mental arithmetic involve the same area of the sensorimotor cortex, and disruption of the shared area leads to both acalculia and finger agnosia [10]. Under the neural reuse view, this co-activation might arise because the shared circuit is specialized for sequencing information or representing arrays, rather than because mental arithmetic is grounded in counting on one's fingers [9]. The neural reuse theory does not deny that action can influence thought, but the theory stresses that not all thought is necessarily grounded in action. The disagreement is one of degree, not of kind.

(b) *How does action experience affect action representation?*

The theory of embodied cognition maintains that the processing or representation of particular actions relies on prior experience doing those actions. In this view, action representation grows out of action experience. The theory predicts that experience with an action should affect subsequent thought relating to that action. Research shows that experience performing an action can influence thought about that action in at least three different ways. It can affect (i) perception of the action, (ii) discrimination of the action, and (iii) comprehension of language related to the action.

Experience performing an action can change how that action is processed in the brain when it is observed (i.e. how the action is perceived). Studies using functional magnetic resonance imaging (fMRI) found that when experts in one dance style were shown a video of that style, areas of their brains involved in action observation and production showed greater activation than when they watched a video of an unfamiliar dance [11,12]. By testing male and female ballet dancers who perform different moves but are familiar with the moves of their partner, a follow-up study demonstrated that it was the dancers' experience of doing an action, rather than their experience of watching their partner perform an action, that accounted for the greater neural activation [13]. These studies suggest that the neural systems involved in action *production* influence the neural systems involved in action *perception*; specifically, having previous experience performing an action is correlated with activation of sensorimotor brain regions when observing that action. Previous motor experience also influences memories of items or objects that we have encountered in the past and the degree to which we like the objects in question [8,14,15].

The kinetic experience of performing an action can help people identify that action even if they have not seen it performed. Casile & Giese [16] blindfolded participants and taught them to swing their arms as if walking using an atypical gait pattern (one with a phase difference of 270° between the left and right arms rather than the typical 180°). Participants who had the kinetic experience of performing the arm motions corresponding to the atypical gait were more

successful at visually discriminating videos of unfamiliar gaits that had phase differences similar to the one they had experienced in training than were participants who had not received the training. However, they were not better at discriminating unfamiliar gaits with other phase differences. These results indicate that specific experience of performing an action improves the ability to visually recognize that particular action—even when the person has never seen the action before.

Previous experience of performing an action can also affect how language related to the action is understood and processed. Using fMRI, Beilock *et al.* [17] measured the comprehension and processing of language related either to ice-hockey movements or to everyday actions. Half of the participants had extensive experience playing ice hockey; the other half had none. The authors found that both groups showed similar comprehension and processing of language related to common actions. However, the group with hockey experience showed greater comprehension of language related to hockey moves than the group without hockey experience. Importantly, the relation between experience and comprehension was mediated by neural activation in the dorsal premotor cortex (believed to be responsible for the selection and planning of well-learned motor sequences [18,19]). The more hockey experience individuals had, the greater their level of activation in this area, and thus the greater their comprehension of hockey-related language. This finding demonstrates that when people hear language related to actions they have previously performed, brain regions involved in planning those actions are activated, which may help them process the language faster and interpret the meaning more accurately than individuals who have not had experience performing the actions. The findings also support previous studies that point to the importance of the left dorsal premotor cortex in auditory comprehension of language related to familiar actions [18–22].

Taken together, the studies outlined above provide support for the embodied cognition framework—namely that the internal representations used to perceive, discriminate and comprehend action and action-based language are associated with the sensorimotor system used to perform these actions [6]. Greater experience performing a certain action strengthens the recruitment of the sensorimotor system in internal representations of information about this action—even in the absence of the overt action itself.

3. GESTURE AND THOUGHT

Gesture forms an integrated system with speech and contributes to the meaning listeners glean from speech [23–25]. For example, listeners are *more* likely to grasp the message conveyed in speech if it is accompanied by a gesture conveying the same message as speech than if it is accompanied by no gesture at all. Conversely, listeners are *less* likely to grasp the message conveyed in speech if it is accompanied by a gesture conveying a different message than if it is accompanied by no gesture at all [26,27]. But gesture goes beyond modulating the listener's comprehension of speech—it can

convey information on its own. For example, listeners can extract information from gesture even if that information is not found anywhere in the accompanying speech [26,27]. Not surprisingly, since gesture forms an integrated system with the speech it accompanies, gestures produced in the context of speech are often difficult to interpret when presented in an experimental situation without speech [28].

There is considerable evidence that gesture plays a role for the speaker as well as for the listener—that it has cognitive as well as communicative functions. Speakers gesture when their listeners cannot see their gestures (e.g. on the phone or when speaking to a person behind a barrier over an intercom [28,29]). More strikingly, congenitally blind speakers (who have never seen anyone move their hands when they talk) gesture and do so even when addressing blind listeners [30]. Findings such as these indicate that gesturing serves a function not only for listeners, but also for speakers themselves. Indeed, speakers are more fluent, producing fewer errors and verbal hesitations, when they are permitted to gesture than when they are prevented from gesturing [31,32]. Gesturing while speaking also frees up working memory: speakers find it easier to remember a list of unrelated items when they gesture while talking than when they do not gesture [33–35]. Gesturing also provides kinaesthetic and visual feedback that can directly aid problem-solving. People can use gesture to work through different solutions to a problem and gather information about the alternatives through the visual and motor feedback of their own gestures [36].

(a) *What do human gestures look like?*

Gestures take many forms. They can be performed with the hands, head or other parts of the body, direct attention towards or away from the speaker and have culturally shared forms or vary according to the speaker's representations. For example, *deictic* gestures draw attention to objects, people or locations in the environment (e.g. pointing at an object or holding it up for display). *Conventional* gestures (or 'emblems') use a standardized form to convey a culturally specific meaning (e.g. an upward movement of the head used to mean *no* in Turkey). *Representational* gestures capture aspects of an action, object or idea either iconically (e.g. moving two fists in the air as though beating a drum) or metaphorically (moving two open hands in the air as though weighing two sides of an argument). These representational gestures are generated on the spot rather than stored in a lexicon (as conventional gestures are), and convey information about a gesturer's thought process or mental representation of an event [25,37].

Representational gestures that depict actions or objects through an iconic mapping to real-world events may be performed from either a first- or third-person perspective. Gestures performed from a first-person perspective are referred to as *character-viewpoint* gestures [25]. In these gestures, the gesturer assumes the role of the person performing the action and his hands represent the character's hands—for example, swinging a closed fist as if gripping the handle of a tennis racket as the gesturer describes a stroke he made when playing tennis.

Gestures performed from a third-person perspective are referred to as *observer-viewpoint* gestures [25]. In these gestures, the gesturer does not assume a role in the action but views it from the outside; his hands then represent participants and objects in the event—for example, tracing the path of a tennis ball as he describes hitting it over a fence.

Not all representational gestures depict aspects of real-world physical events. Representational gestures can also be used to represent abstract ideas. When they do, they are usually described as *metaphoric* because they map abstract ideas onto physical actions or features. The gestures themselves are not metaphoric; they convey physical features, movement or space. Rather, the metaphor is contained in the relation between gesture and speech, where speech communicates an abstract concept and gesture adds a physical element to the concept, often providing a link to an action that grounds the abstract language in physical experience. For example, a person might say, 'we need to think about the future' and extend a hand forward, thereby displaying a temporal metaphor in which the future is ahead of the speaker. In one common type of metaphoric gesture, the speaker gestures as if holding a solid object in one or more hands while talking about an abstract concept or idea. By gesturing as if holding onto the idea, the speaker indicates that she is treating the idea as a physical object, as though it were a thing that can be given from one person to another, lost, taken apart, etc. Metaphoric gestures can also convey abstract relations by emphasizing parts of the accompanying speech or surrounding physical environment. In one study of mathematical problem-solving, children indicated that the two sides of an equation should receive equal treatment by producing the same sweeping motion under each side of a mathematical equation during their explanations [38]. Although the children's gestures did not convey traditional metaphors, they did highlight an abstract relation (the notion of equivalence) by gesturing to each side of the written equation in exactly the same way. Examples such as these demonstrate how gesture can ground even abstract ideas in physical actions.

(b) Gesture can represent action

Representational gestures are thought to be a type of simulated action (e.g. [25,39]). Recently, Hostetter & Alibali [40] proposed that these gestures result from a direct extension of mentally simulated action and perception. In their view, gesture arises when activation spreads from the areas involved in action planning to those involved in action execution. Character-viewpoint gestures provide support for the view that gesture is rooted in action simply because they resemble the kinematics of actions on objects in the real world.

In character-viewpoint gestures, the actions of the gesturer's hands closely mimic the movements she would make when performing the action in the real world. This similarity may be used, in certain circumstances, to enact familiar action sequences while reasoning or talking about real-world action. The proprioception of performing familiar movements may activate detailed mental representations of objects by simulating acting on the world [41]. Streeck describes

the gestures a car mechanic made while talking about problems with different cars. Because the mechanic frequently encountered the same types of problems, he had developed a set of 'habitualized' gestures he used when faced with familiar problems. These gestures had similar forms every time he used them (such as turning an invisible ignition key or shifting an imaginary car into second gear) and were closely based on the motor patterns he used when solving the problems in the real world. These types of routinized gestures lie somewhere between iconic representational gestures and conventional gestures because they use the same movement pattern every time.

Peoples' gestures tend to reflect their own experience. For example, Cook & Tanenhaus [42] found that the gestures people produced when talking about a particular task (the Tower of Hanoi puzzle) reflected their kinematic experience solving the problem. The Tower of Hanoi is a challenging task in which people are presented with an array of three pegs in a row and are asked to move a stack of discs of different sizes from one peg to another in a particular order so that larger discs are never placed upon smaller ones and only one disc is moved at a time [43]. Cook & Tanenhaus [42] had adults solve the task and then asked them to explain how they had solved it. When people completed the task on an actual tower before describing how they solved it, they used many character-viewpoint gestures in their descriptions: cupping and moving their hands as if holding and transferring discs up and over the peg. When people solved the task on a computer, they produced fewer grasping handshapes during their descriptions and the trajectory of their gestures was more likely to reflect the horizontal path that the mouse followed (i.e. they moved horizontally from peg to peg rather than moving up and over each peg). Gestures representing actions on or by objects thus reflect the speaker's previous experience with those objects.

Gestures representing the use of objects in actions (as in tool use) are cognitively complex. They require that the gesturer represent not only the motion of the action (say hammering) but also the object involved in the action (the hammer). To depict the use of an object, a gesturer must either gesture as if holding an imaginary object, or use a body part (usually the hand or finger) to represent the object involved in the action. Imaginary object gestures are a type of 'character-viewpoint' gesture because the hands are representing the hands of the agent holding the object. Gestures in which a body part represents an object are 'observer-viewpoint' gestures because the gesturer does not act as an agent manipulating an object but instead depicts only the action of the object. Imaginary object gestures more closely resemble the actions made when acting on real-world tool objects (e.g. moving the hand shaped as though holding a toothbrush back and forth across the mouth when describing brushing one's teeth). However, they require that the gesturer have a strong mental representation of the tool object involved in the action because there is no physical placeholder standing in for the tool. In contrast, gestures in which a body part represents an object rely on physical substitutes for the object involved in the action (e.g. rubbing the index finger across one's teeth during a description

of tooth brushing) and thus might require a less strong or detailed mental representation of the tool object than imaginary object gestures.

Gestures depicting tool use have not been studied in spontaneous conversation, but there is an extensive experimental literature on the types of manual representations people produce when asked to pantomime how tools are used. Adults pantomime these types of events as if holding an imaginary tool in their hand most of the time [44]. But 3- and 4-year-old children frequently use body parts as stand-ins for the tool object rather than manipulating an invisible tool [44–46]. For example, they run their fingers through their hair when asked to portray a hair-combing act, rather than pretending to hold a comb and move it over their hair. One possibility is that the children are not using their hands to *represent* action, but are instead merely performing the act with their fingers (i.e. literally combing their hair with their fingers). The same phenomenon has been found in aphasics [47] and schizophrenics [48], individuals whose symbolic representation systems have been disrupted.

It is unclear whether pantomimes elicited to portray tool use are cognitively different from gestures spontaneously produced to communicate about tool use. Some have argued that tool-use pantomimes involve different neural substrates from those involved in producing communicative gestures (see review in [49]), a distinction supported by the fact that apraxic patients who have difficulty producing tool-use pantomimes have fewer (or no) problems producing conventional gestures or meaningless hand shapes [50,51]. In contrast, Frey [49], who finds no differences in activation during tool-use pantomimes and communicative gestures, argues that difficulty producing tool-use pantomime is due to the cognitive demands of representing absent objects.

Gestures in which hands represent hands (and act upon imaginary objects) intuitively seem less cognitively complex than those in which hands represent other things. However, it is clear from the research on tool-use pantomimes that manipulating imaginary objects in gesture is a non-trivial task and involves more than simply recreating the motor patterns performed during actions on the real world. Mental representation of non-present objects is difficult, and people with incomplete linguistic representation systems often rely on a body part to stand in for the absent object. These difficulties highlight the difference between performing a movement sequence as part of a real-world action and performing the same sequence as part of a representational gesture. The kinetic movements may be very similar, but using movement to *represent* action adds an additional level of complexity.

Gesturing from a first-person perspective (as in imaginary object gesture) may be complex not only because the gesturer needs to mentally represent an imaginary object, but also because the gesturer needs to take the perspective of the agent in the depicted event. Character-viewpoint gestures as a whole seem to involve a more sophisticated mental representation of events than observer-viewpoint gestures because of the need to take the agent's perspective into account. This perspective-taking ability is associated with narrative development in speech. Young children's ability to produce character-viewpoint gestures is associated with better concurrent narrative skills and predicts improvements in narrative skill in the future [52].

(c) Gesture can influence thought

A great deal of research has shown that the spontaneous gestures speakers produce provide a window onto their thoughts (see [37] for a review). But there is growing evidence that gesturing can go beyond reflecting thought and can play a role in changing thought. In order to demonstrate that gesturing is causally involved in thinking, we need to manipulate the gestures that speakers produce.

Broaders *et al.* [53] asked children to gesture while explaining their solution to a math problem and subsequently gave them instruction on the problems. Children who were asked to gesture before the lesson were more likely to benefit from the subsequent lesson than children who were not asked to gesture. Many of the children conveyed strategies in their gestures that they had not expressed before being asked to gesture. Being forced to gesture activated previously unexpressed concepts. In turn, this expanded repertoire led the children to profit from subsequent instruction.

Gesturing can convey cognitive benefits to the speaker even when speakers are told precisely how to move their hands. Ehrlich *et al.* [54] gave a mental-rotation task to two groups of children, each instructed to gesture in a different way. In the task, children were shown two unconnected shapes and were asked to choose from an array of images the shape the two separated pieces would make if they were moved together. The unconnected shapes needed to be moved horizontally or vertically or rotated to create the final shapes. During a mental-rotation lesson, one group was told to show the experimenter with their hands how they would move two pieces together. Children in this group produced both character-viewpoint gestures (e.g. rotating their hands as if moving the pieces) and observer-viewpoint gestures (e.g. tracing the trajectory that the pieces would take). The other group was told to point to the two pieces. The children who produced gestures (either character- or observer-viewpoint gestures) representing the movement of the pieces learned more from the mental-rotation lesson than did children who produced pointing gestures [54].

As another example, Goldin-Meadow *et al.* [55] taught children to gesture in a particular way during a lesson on mathematical equivalence. The gestures (in which a pair of numbers was grouped together by placing a 'V' handshape underneath them) conveyed a novel 'grouping' strategy that none of the children had used before. The children were then given a lesson on mathematical equivalence and were told to perform the gestures they had learned. Importantly, the new grouping strategy was never used by the teacher, in either gesture or speech. Children who had been told to gesture using the grouping strategy improved more from the lesson than children who were not told to gesture. Moreover, the children who improved were very likely to express the grouping strategy in speech on the

post-test, even though they had never heard it expressed in speech during the lesson. Gestures can thus instil new ideas in learners—creating thought in addition to reflecting it.

4. GESTURE AS A BRIDGE BETWEEN ACTION AND THOUGHT

Both action and gesture can affect the mental representation of actions and objects, but gesture's ability to *represent* action offers a way to ground abstract ideas in concrete actions. Gestures that represent action are actions performed within an imagined world. When gestures simulate action on or by objects, the objects involved in the event must be represented mentally. Actions, on the other hand, are performed on the physical environment. The objects they act on are present and do not need to be represented mentally. Thus, when we perform actions on objects, we are able to offload some properties of the task onto the environment. However, when we use gesture to represent action on or by objects, we must rely on mental sensorimotor representations of the objects involved. This is particularly true for gestures in which the gesturer's hands manipulate imaginary objects. Gestures in which a body part is used to represent an object involved in an action are symbolically complicated because they use one thing (e.g. a hand) to stand for another thing (e.g. a toothbrush), but they also allow some cognitive offloading because the hand serves as a physical placeholder for the object.

Simulating an action on an imagined object in gesture seems to strengthen the link between the action and the mental representation of the object, and does so *more* than performing the action on the object in the physical world. Beilock & Goldin-Meadow [56] asked participants to solve the Tower of Hanoi puzzle with real discs, and then describe how they solved the puzzle to another person. The largest disc was on the bottom of the stack and needed two hands to lift; the smallest disc was on the top and could be lifted with one hand. Following their explanation, participants were divided into two groups and given the task again. One group solved the task with precisely the same discs (*no-switch*); the other group solved the task with discs whose weights had been reversed (*switch*)—now the smallest disc on the top was the heaviest and needed two hands to lift. Participants in the *switch* group who had gestured with one hand when describing moving the smallest (and lightest) disc found it harder to execute the task the second time than the first. Moreover, their performance on the reverse weight task was predicted by the number of one-handed gestures they made during their explanation of the first task—the more one-handed gestures they produced, the worse they did on the task when the weights were reversed and the smallest disc required two hands to lift. The gestures produced by participants in the *no-switch* group had no relation to their performance on the second task. These findings suggest that people who used one-handed gestures to represent moving the small disc represented the disc as light, even though weight was not a relevant factor in solving the task. Representing the small disc as light causes problems when solving the problem a second time in the *switch* group (where the small disc is no longer light), but not in the *no-switch* group (where the small disc is still the lightest). Importantly, when additional participants were given the same tasks but without the explanation phase in between, there were no differences between the *no-switch* and *switch* groups—that is, switching the weight of the discs had an effect on subsequent performance *only* when the participants gestured prior to the performance, and *only* when those gestures were incompatible with the performance. In a follow-up study, Goldin-Meadow & Beilock [57] found that gesturing about the task more strongly influenced mental representations of the actions involved in the task than performing the task again (i.e. than acting on the objects). These studies add weight to the claim that representing action in gesture embeds embodied information into mental representations of action. In fact, when the effects of gesturing about action and acting were pitted against one another, gesturing appeared to have a stronger effect on the mental representation of the action than performing the action itself had.

5. ACTION AND GESTURE IN NON-HUMAN PRIMATES

Non-human primates (specifically simians, hereafter referred to as 'primates') are extremely adept at performing manual actions. Although not as dexterous as humans, primates are nonetheless able to execute a great number of manual tasks requiring fine motor control (e.g. extractive foraging, delicate grooming and tool use). They are also able to extract information (including how to accomplish certain tasks) from watching others perform actions. Moreover, some primate species, great apes in particular, use a wide range of gestures in communication. Their gestures are used flexibly and intentionally, and at least some communicate specific meanings (see review in [58]). However, the gestures that primates produce lack the representational elements of human gesture. Whereas many human gestures *symbolize* actions and objects, ape gestures primarily *indicate* the gesturer's future actions by performing an abbreviated part of the action that would, in its full version, fulfil his or her goal.

(a) *What do primates know about actions?*

Primates are able to recognize particular movements in themselves and to determine when their movements are the same as those of others. They can easily learn to perform new actions. However, they are more likely to focus on the goal and the primary method of an action than on the details of specific movements used to achieve the goal [59]. This observation has led some (e.g. [60]) to consider primates 'emulators' and human children 'imitators', although meta-analysis across studies shows that primates are capable of both goal emulation and process imitation (see [59,61] for discussion). Even though primates tend to focus on obtaining desirable outcomes rather than on a specific means for achieving those outcomes, they are able to detect small details in movement. For example, some ape species can recognize when an experimenter is copying their movements exactly

[62] and, like humans, apes that are being copied will sometimes try to trick the copying individual into performing bizarre actions or making a mistake. The ability to recognize and learn both the kinematics and goals of actions from others suggests that primates have a mental representation of what action needs to be performed on a particular object, and can form or modify that representation without acting on the object themselves. There is no evidence, however, that primates are able to *actively* manipulate their representations and rehearse their actions before they attempt an action.

When solving unfamiliar tasks, primates are able to modify their techniques and strategies in response to information acquired during trial-and-error learning, but there is little evidence that they reason through multiple solutions to a problem (so-called 'mental rehearsal') before undertaking any actions ([63]; but see [64]). It is, of course, impossible to say exactly what is going on inside the minds of non-human primates during problem-solving, but they do not exhibit the external behaviours that are associated with mental rehearsal in humans, such as gesturing or practising actions out of their functional contexts.

Early studies with great apes (e.g. [65]) suggested that primates might, indeed, consider different possible outcomes when faced with difficult problems, but there has been no consistent evidence of primates either gesturing through or acting out different versions of their actions before they act. Kendon [66] notes that several chimpanzees tested by Köhler [65] on problem-solving tasks behaved as though they were 'acting out the wished-for state of affairs in a situation that [they treated] as analogous to the actual one' ([66], p. 210). In the examples Köhler and Kendon describe, chimpanzees were presented with a challenging task (stacking boxes, uncoiling a rope, lifting a cage) that they had to perform to gain access to a food item. Köhler describes several cases in which an ape, when confronted with a problem, performed non-functional actions or hand movements that were not directed towards solving the problem at hand. These actions or gestures were thought to be an indication of 'working through' the problem before attempting a solution. In one such case, a chimpanzee was presented with a room full of boxes and a banana suspended in the air out of reach. The chimpanzee moved one box underneath the fruit and then eyed the distance from the box to the banana. Then he retrieved a second box, 'but, instead of placing it on top of the first, as might seem obvious, began to gesticulate with it in a strange, confused, and apparently quite inexplicable manner. He put it beside the first, then in the air diagonally above, and so forth' ([65], pp. 46–47). Kendon argues that the aborted actions are 'pre-enactments' of different scenarios, and that the chimpanzee 'embarks on a course of action with the second box, but each time foresees that its outcome will not suit his purposes, so he cuts off, changes course and tries again' ([66], p. 210).

These examples are intriguing, but they were made as real-time observations during problem-solving tasks and there have been no comparable observations since that would allow us to conclude that apes do use action

or gesture to plan their actions. The more common view is that apes do not rehearse actions. In fact, their lack of rehearsal has been used as evidence that primates are incapable of 'mental time travel' (i.e. imagining performing actions in the past or the future [67]). Primates, like many animals, do perform modified versions of actions (such as biting, fighting or courtship) during play interactions [68,69]. However, while behaviours 'rehearsed' during play may help young primates perfect adult behaviours and learn to negotiate social situations, they differ from the targeted mental rehearsal involved in thinking through different versions of an action during action planning.

(b) Primate gesture

It is difficult to directly compare reports of gestures in humans and primates because researchers working in the two areas define gestures differently and often address different questions. Researchers studying primates define gesture as including not only visual movements of the hand, face and body (*visual* gestures), but also movements that come into contact with other individuals (*tactile* gestures) and movements that produce audible sounds (*audible* gestures). Primate researchers are also more likely to focus on gestures that are directed towards other individuals and discount similar movements made when an animal is solitary. These decisions make it particularly hard to determine whether primates ever use gesture as a cognitive aid outside of communicative contexts. The communicative gestures primates produce can be directly compared with communicative gestures produced by humans, although the challenges of determining when a gesture is intentionally communicative and what the gesturer aims to communicate are more difficult when observing primates.

Like humans, primates frequently use facial, manual or whole-body signals in communication. But primate gestures differ considerably from human gestures, particularly when it comes to symbolic representation of the world. Many non-vocal primate signals appear to be involuntary responses to internal emotional states like fear or excitation [58]. Involuntary signals can be effective in communicating the presence of recurring events or goals (e.g. signalling the presence of a predator or asking to mate), but they cannot be employed strategically and almost certainly do not provide a cognitive aid in the way that human gesture does. However, some types of primate gestures are used flexibly in communication. These gestures, observed predominantly in great apes, are often referred to as 'intentional gestures' [70,71].

All great apes gesture to communicate. A large cross-species comparative study of great ape gesture found more similarities than differences in the types and uses of gesture across species ([71]; see also [72–75]). Each species had a comparable repertoire size of 20–30 visual and tactile gestures. Subsequent studies have recorded species repertoires of up to 100 gesture types (e.g. [76]), but these differences can largely be attributed to how narrowly each gesture type is defined in each study [77].

Great apes use gesture in purposeful and socially complex ways. Their choice of when and how to gesture,

particularly their choice of the tactile versus visual modality (whether they touch or do not touch others), is sensitive to whether they can be seen by others (e.g. [71,76,78–80]). This finding is thought to be evidence that apes take the visual attention of others into account when choosing how or when to gesture. There is also evidence that apes gesture to achieve particular goals. They expect responses from others and often wait for a response if a recipient does not respond immediately (see results from several species in Call & Tomasello [71]). Moreover, at least some ape gestures have specific meanings, and apes often repeat, change or elaborate their gestures when a recipient responds in an undesired way [70]. When a recipient fails to respond at all to an attempted gesture, apes will persist and elaborate their gestures according to whether or not the recipient can see them [81]. There is also some evidence that apes (at least orangutans) tailor their communicative strategies to how successful their initial communicative attempt was, so that they repeat gestures more often when communication has been partially successful and use a wide range of gestures when communication has failed completely [82].

(c) Comparison to human gesture

Though apes display a sensitivity and flexibility in their gestures that indicate they can use gesture to communicate intentionally, their gestures are distinctly different from the gestures used by humans. Human and ape gestures differ in the degree to which they are combined in *structured* ways, whether they can communicate a wide range of *meanings*, and whether they *represent* or reference real-world events in the same way.

(i) Structure

Apes can combine gestures with one another, and two apes can produce gestures in response to one another in a communicative exchange. However, there is no indication that these sequences of gestures are combined according to systematic patterns—either to attract attention before communicating a particular desire, or to communicate a more complex meaning than is possible using a single gesture. Apes' gesture combinations are typically either repetitions of the same gesture or different types of gestures with the same meaning [83,84].

Humans rarely combine the spontaneous gestures that they produce along with speech into gesture sequences [25]. However, when humans gesture without vocalizing (which is the typical situation for apes, who rarely produce their gestures along with vocalizations), they not only routinely combine different manual gestures with one another, but they do so following a systematic pattern; in other words, they use devices characteristic of human language. The clearest example is the sign languages of deaf communities handed down from generation to generation (e.g. [85]). However, hearing individuals will also develop complex patterns of gestures when they interact in circumstances where speech is either impossible or inappropriate (e.g. workers exposed to high noise levels or people following religious conventions prohibiting speech), although these systems rarely achieve the complexity characteristic of sign languages. Kendon notes that 'the more *generalized* [the] circumstances are, the more complex

[the] systems become' ([86], p. 292). Thus, systems restricted to a specific type of interaction—say, operating heavy machinery—do not face pressures to adopt greater complexity because they are not used frequently enough, and in enough different scenarios to require significant modification. When human gesture systems are used frequently in a variety of situations—as in the case of the sign languages of the Plains Indians and Australian Aborigines—they begin to take on the complexities of spoken language.

Strikingly, humans will combine gestures in language-like ways even when they have never been exposed to the structures of a conventional language (spoken or signed). For example, deaf children whose hearing losses prevent them from acquiring the spoken language that surrounds them, and whose hearing parents have not exposed them to a conventional sign language, invent gestures to communicate with the hearing individuals in their worlds. These gestures exhibit many of the properties found in human language, including a simple syntax based on gesture order [87–91]. As another example, when hearing adults with no knowledge of sign language are told to describe a series of events using only their hands, the sequences of gestures they produce tend to follow a systematic order [92,93]. Interestingly, all hearing adults tested in this way display the same order (subject–object–verb), an order that is found in half of the world's spoken languages, and they do so whether or not the order is predominant in their spoken language [94–97].

(ii) Meaning

In contrast to human gestures, ape gestures are almost universally requests for a particular response from the recipient. Tomasello & Camaioni [98] used this observation to draw a sharp contrast between ape and human gestures, characterizing apes' gestures as *imperative* and children's gestures as *declarative*. Where humans (even infants) will gesture to draw attention to an object or to comment on an aspect of the world, apes gesture primarily to request others to interact or leave. Most of the gestures of one ape genus (orangutans) can be categorized into only six types of requests: affiliate/play, move away, share food/object, stop action, co-locomote and take food/object [70]. Other ape species use gesture to communicate fairly similar meanings (e.g. [76]). So, whereas human gestures can communicate a potentially boundless number of meanings, primate gestures appear to be restricted to initiating, ending and moderating frequent kinds of social interactions.

(iii) Representation

Another striking difference between ape and human gesture is the lack of deictic and representational elements in apes. Humans use deictic gestures to draw attention to objects in the environment and representational gestures (i.e. character-viewpoint, observer-viewpoint and metaphoric gestures) to refer to objects or events. Great apes almost never use gestures deictically to draw attention to things in conspecific interactions, although deictic gestures are sometimes used by captive apes communicating with humans ([99,100]; for an example of deixis in the wild, see Pika & Mitani [101]). Most importantly, the gestures

apes use in their natural communication systems, even when produced to communicate with humans, do not seem to have any of the representational elements found in human gestures. Many gestures appear to be an incipient action reduced from a full-blown action that evoked a particular response from a recipient in the past; a process called *ontogenetic ritualization* [71,98]. For example, a shoving action eventually becomes the gesture 'nudge' or 'shoo' as the recipient learns to predict the gesturer's behaviour from the start of the action and responds appropriately.

It seems likely that most ape gestures began as actions and were co-opted into communicative devices either during ontogeny (via ritualization) or over evolutionary time [102]. Indeed, even the few ape gestures that have been reported to be iconic could have been ritualized from functional actions rather than representing actions in the same way that human gesture does. One commonly cited 'iconic' gesture involves an ape's gesturing to indicate the direction it would like another to move by brushing along the recipient's body or swinging an arm in the desired direction (e.g. [103–105]). It is possible that these gestures indicate the direction of desired movement through iconic representation of the action. But they also may be incipient actions or other movements ritualized into gestures from what were once effective pulling or guiding actions. If the latter is the case, then the similarity between the movement of the gesture and the desired action would be incidental rather than an iconic representation of action.

(d) *Could gesture serve as a bridge to thought in primates?*

It is clear that humans gesture not only to communicate but also to aid their own cognition. The fact that humans gesture to themselves (outside of communicative contexts or when they cannot be seen) has been taken as evidence of gesture's cognitive function. Unlike humans, naturally reared apes have not been found to gesture when alone or when they are behind a barrier. In fact, apes choose not to use manual gestures when they cannot be seen and instead switch to vocal signals or auditory gestures (e.g. [79,81]). One methodological difficulty in making this comparison between apes and humans is that most studies of ape gesture require that a manual movement be directed towards another individual in order for it to be counted as a gesture [102]—in other words, if an ape were to produce a gesture-like movement in the absence of a partner, it would not meet one of the criteria for a gesture.

Although we cannot conclude that the gestures apes use have no effect on their cognition, it seems safe to say that their gestures do not contribute to building mental representations the way humans' gestures can. There is no evidence that apes use gesture in a truly *representational* way. Their action-like gestures 'represent' actions through learned association, not by design. Many, if not most, of the gestures apes use are 'species typical' and do not differ across individuals or groups [70,71,76]. Apes can use their gestures flexibly in response to social and communicative contexts (varying when and where they gesture and which gestures to use), but the underlying forms of their gestures are probably chosen from a pre-existing repertoire. This tendency to use the same gestural forms every time differs sharply from humans' use of representational gestures, in which the exact forms are spontaneously generated during communication; they are not emblems or lexical forms—they have no 'wrong' forms.

Primates are excellent observers of actions and signals. They can extract information about the world by learning relationships between the signals and subsequent actions of other individuals or between others' signals and events in the external environment [106]. They can understand, interpret and predict actions of others, even when those actions occur out of view [107]. They are able to learn complex novel actions through observing others [108], and this ability to socially learn manual actions probably contributes to 'cultural traditions' in food processing or manipulation of objects [109,110]. However, even though primates can process, learn from and replicate actions, there is no evidence that they can represent actions using gesture.

When placed in the right environment, apes can acquire symbolic communication systems, learning the associations between objects in the world and symbols representing those objects. If apes are taught human-designed communication systems (such as modified American Sign Language or computer-based symbols), their communication resembles, in some but not all respects, the communication of a 2-year-old child (e.g. [111–114]). Moreover, there is some indication that when apes are given access to a symbolic representational system, they can use the system for more than communication. For example, one of the most proficient ape signers, Washoe, used her signs appropriately when she was alone, signing 'quiet' when sneaking into a room or signing to her dolls [115]. However, the vast majority of the communication that language-trained apes produce not only is directed towards another, but is also used to get that individual to do something (i.e. to make a request [116]).

We do not claim that primates are incapable of mentally representing actions or objects, but it is clear that they do not represent actions or objects in their gestures. Without representational gestures, primates cannot link action to mental representation in the same way humans do. It is noteworthy, however, that when primates are taught a symbolic communication system, they do at times exhibit behaviours—such as using communicative signals outside of communicative contexts—that suggest they may be able to use symbols to aid or complement cognition (see, for example, [117,118]). Language-trained apes provide an interesting comparison to both humans and non-language-trained apes because they demonstrate the level of abstract cognition apes can reach when reared in human-like conditions and highlight the importance of rearing environment in the development of cognitive and communicative abilities.

6. GESTURE AND MENTAL REPRESENTATION IN THE EVOLUTION OF LANGUAGE

The gestures primates use in their natural communication systems have only little in common with the

types of human gestures we have discussed in this paper (deictic, conventional and representational). However, they do resemble human language more than primate vocalizations do. Unlike humans, primates cannot learn new vocalizations (their vocal repertoires are essentially fixed) and their vocalizations seem to be elicited by emotional states rather than employed intentionally to communicate particular goals [119]. Primates have greater flexibility and control in manual communication than they do in vocal communication. They can easily learn new hand movements, and use gestures flexibly in response to the visual attention and reactions of others. In fact, this flexibility in primate gesture is often cited as support for the theory that human language originated as a gestural system.

Many have proposed that human linguistic structure first emerged in gesture and only later spread to vocalization (e.g. [66,120–123]). The prevalence of co-speech gesture in human language [25] and findings that gesture precedes and predicts children's development of spoken language [124,125] demonstrate that gesture is an integral part of modern human language and not something layered on top of an older verbal system. Representational gesture, in particular, has been suggested as having provided a means of communicating complex events before human ancestors developed the ability to use shared symbols [121]. Indeed, some argue that representational gesture (or, rather, pantomime) was a critical stage in a progression from manual action to spoken language and propose the mirror neuron system as a neural foundation for this transition [126].

(a) Mirror neurons

The discovery of mirror neurons provides a possible device through which primates might identify similarities between their own movements and the movements others produce. Mirror neurons are visuomotor neurons found in area F5 (and other connected areas) of the primate premotor cortex (roughly analogous to Broca's area in humans). They are unusual in that they discharge both when a primate performs an action directed towards an object and when it watches another individual perform that same action [127–129]. These neurons provide a link between perceived and performed actions and are one possible mechanism through which observed action could become *simulated* action.

The majority of work on primate mirror neurons has been done on macaque monkeys using single-cell recording techniques. These studies have found several different types of mirror neurons distributed in different areas of the brain. Some neurons respond primarily to the goals of actions (e.g. picking up an object), whereas others respond to both the goals and specifics of the movements (e.g. picking up an object between two fingers) [130]. Primate mirror neurons are primarily activated by the movements or goals of grasping, placing or manipulating actions, and most are specific to one of these actions (i.e. they are activated by only one type of action [131]). Importantly, however, most monkey mirror neurons respond only when these actions are directed towards

physical objects; they do not recognize movements (such as gesture) that simulate goal-directed actions in the absence of those objects [107]. Monkeys do not have to see the object to activate the mirror system, but they do have to 'believe' that the object is present. For example, if a monkey is first shown an object and the object is then hidden by a screen, the monkey's mirror neurons will fire when a grasping hand reaches for the now-hidden object (although the response will be smaller than when the object is visible [132]). If the monkey is shown the grasping hand reaching towards a screen without having first seen the object behind the screen, the monkey's mirror neurons will not fire [132]. Thus, if a grasping movement is directed towards an empty space (as it would be during a representational gesture), primate mirror neurons will either not fire or produce only a weak signal [131,133].[1] Interestingly, primate mirror neurons respond to the sounds made by manipulating specific objects (e.g. ripping a piece of paper [135]) though neither the action nor the object is visible. This strengthens the argument that is it the 'belief' that an object is present, rather than the physical presence of the object, that activates the mirror system.

Evidence for a mirror neuron system in humans comes primarily from imaging studies (such as fMRI) and techniques stimulating areas of the brain during behavioural tests (such as transcranial magnetic stimulation) [107,130]. The human mirror system appears to have many of the same properties as the monkey mirror system. It fires for specific motor patterns as well as the goals of motor acts [130]. However, the monkey and human mirror neuron systems differ in at least one critical respect: the monkey mirror neuron system does *not* fire unless an object is present (or the monkey thinks the object is present); the human mirror neuron system does. The human mirror system responds to empty-handed gestures, that is, to movements made in the air, simulating actions made on an object but without having the object present (though the brain areas that respond to representational actions are not entirely the same as those that respond to object-directed actions [136,137]; see also Skipper *et al.* [138], who find activation of the human mirror system during processing of co-speech gestures). This neural response to simulated action in the absence of objects may provide the foundation for understanding gestures as representations of actions on or by objects. The important point from the point of view of our discussion here is that this type of neural response is found in humans but not in monkeys.[2]

(b) A cognitive leap?

Arbib [139] proposes that the ability afforded by the mirror neuron system to draw parallels between actions of the self and others paves the way for complex imitation and provides a foundation for the evolution of neural mechanisms supporting representation through pantomime. Pantomime, he argues, was a necessary precursor to *protosign*, which when combined with vocalizations, evolved into *protolanguage* in the human lineage. The ability to recognize and imitate the manual actions of others is undoubtedly necessary for

a complex gestural communication system to emerge, but it remains unclear how, when and why human ancestors began using gesture to represent elements of the world around them.

In the wild, primates are never exposed to representational gestures. Though many primate species have a rich repertoire of facial, manual and bodily signals, primate gestures lack the representational elements characteristic of many human gestures. Although some primate gestures have predictable meanings, the meanings are not iconically represented (as many human representational gestures are) nor are they culturally variable (as human conventional gestures are). Moreover, although primates can learn new gestures, they do not seem to acquire them through cultural transmission. The majority of gestures used by particular primate species do not vary greatly among captive populations. Idiosyncratic gestures unique to individual animals are frequently observed, but they do not spread through populations to become group-specific gestures as you would expect if gestures were acquired via cultural transmission (e.g. [73–76]). For primate gestural systems to have developed over time to produce anything resembling pantomime or conventional gestures, primates would have had to develop the ability to add gestures to their communicative repertoires by observing others. There is no evidence that primates have this capacity.

Using manual gesture to simulate and represent actions and objects outside of the context of acting on real-world objects represents a cognitive leap in hominid evolution. It is possible that increased demand for accurate tool manufacture and use in human ancestors drove many cognitive developments, including neural lateralization, more complex mental representations and complex manual imitation. These changes could have altered the nature of the gestural communication system, allowing human ancestors to acquire new gestures through imitation and link them to representations of actions. It seems likely, however, that sweeping social changes would also have been necessary for human language to develop. Studies in which apes are taught human communication systems have demonstrated that apes can learn new gestures or symbols and use them referentially (although not combinatorially), but even when acquiring symbolic communication, apes do not develop the social and representational milestones, such as theory of mind and pretend play, that accompany language development in young children.

The extent to which rearing environment, linguistic development and cognitive development interact with one another is a topic of great interest in human research and is not at all understood in primates [58]. A handful of studies have shown that apes raised in a human-like environment exhibit cognitive skills that apes reared in their natural social groups do not exhibit [140,141]. However, it is not clear which aspects of the rearing environment are most influential and whether the cognitive abilities of human-reared apes are truly different from those of naturally reared apes; apes could differ from humans in their external behaviours because they differ in their motivation to participate in certain types of

activities, motivation provided, in part, by rearing the apes in human environments. Given that juvenile primates (especially humans) have an extremely long period of maturation and dependence [142], there is great potential for interaction between the rearing environment and neural, cognitive and communicative development.

We know that the role gesture plays in human language development is complicated. Gesture both responds to and influences linguistic and environmental variables. Parents' gesturing predicts children's gesturing, which, in turn, precedes and predicts child speech [143]. Children's gesturing also alters the environment for the child by facilitating interaction with parents and thus enriching the child's linguistic input [144]. Comparative developmental research is necessary to investigate and tease apart the respective contributions of environment, action, gesture and cognition to non-human primate communication systems. That said, representational gesture (and the cognitive advantages it brings) appears to be a uniquely human ability. It is unclear, however, which pieces of the puzzle are missing in extant primates. We do not know whether primates lack a neural substrate enabling complex mental rehearsal, the cognitive ability to connect gesture to mental representation or the social motivation to create a rearing environment that would foster the development of these abilities. We hope that future studies on primates will investigate the relationships among action, gesture and cognition during development. Such studies will not only help us understand how these variables influence one another in primates, but also shed light on the relationship among action, gesture and representation in human evolutionary history.

ENDNOTES

[1]The one exception is facial movements that are not directed towards objects (e.g. lip-smacking [134]). Monkey mirror neurons do respond to these acts even though an object is not involved. However, these facial movements, although communicative, are not representational in the way the human manual gestures we discuss here are.

[2]It is important to note that all of the work on the primate mirror system has been done on monkeys, but most of the findings of complexity in gestural communication come from great apes. It is possible that the mirror system in great apes is more human-like than monkey-like; however, given the fact that ape gestures lack representational elements, it seems likely that their mirror systems are still significantly different from those of humans.

REFERENCES

1 Varela, F. J., Thompson, E. & Rosch, E. 1991 *The embodied mind.* Cambridge, MA: MIT Press.
2 Wheeler, M. 2005 *Reconstructing the cognitive world.* Cambridge, MA: MIT Press.
3 Barsalou, L. W. 1999 Perceptual symbol systems. *Behav. Brain Sci.* **22**, 577–660.

4 Garbarini, F. & Adenzato, M. 2004 At the root of embodied cognition: cognitive science meets neurophysiology. *Brain Cogn.* **56**, 100–106. (doi:10.1016/j.bandc.2004.06.003)

5 Glenberg, A. M. 1997 What memory is for. *Behav. Brain Sci.* **20**, 1–55.

6 Wilson, M. 2002 Six views of embodied cognition. *Psychon. Bull. Rev.* **9**, 625–636. (doi:10.3758/BF03196322)

7 Zwaan, R. A. 1999 Embodied cognition, perceptual symbols, and situation models. *Discourse Process.* **28**, 81–88. (doi:10.1080/01638539909545070)

8 Beilock, S. L. & Holt, L. E. 2007 Embodied preference judgments: can likeability be driven by the motor system? *Psychol. Sci.* **18**, 51–57. (doi:10.1111/j.1467-9280.2007.01848.x)

9 Anderson, M. L. 2010 Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* **33**, 245–266. (doi:10.1017/S0140525X10000853)

10 Roux, F.-E., Boetto, S., Sacko, O., Chollet, F. & Tremoulet, M. 2003 Writing, calculating, and finger recognition in the region of the angular gyrus: a cortical stimulation study of Gerstmann syndrome. *J. Neurosurg.* **99**, 716–727. (doi:10.3171/jns.2003.99.4.0716)

11 Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. & Haggard, P. 2005 Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb. Cortex* **15**, 1243–1249. (doi:10.1093/cercor/bhi007)

12 Rizzolatti, G., Fogassi, L. & Gallese, V. 2001 Neurophysiological mechanisms underlying understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)

13 Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E. & Haggard, P. 2006 Seeing of doing? Influence of visual and motor familiarity in action observation. *Curr. Biol.* **16**, 1905–1910.

14 Ping, R., Dhillon, S. & Beilock, S. L. 2009 Reach for what you like: the body's role in shaping preferences. *Emot. Rev.* **1**, 140–150. (doi:10.1177/1754073908100439)

15 Yang, S., Gallo, D. & Beilock, S. L. 2009 Embodied memory judgments: a case of motor fluency. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 1359–1365. (doi:10.1037/a0016547)

16 Casile, A. & Giese, M. A. 2006 Nonvisual motor training influences biological motion perception. *Curr. Biol.* **16**, 69–74. (doi:10.1016/j.cub.2005.10.071)

17 Beilock, S. L., Lyons, I. M., Mattarella-Micke, A., Nusbaum, H. C. & Small, S. L. 2008 Sports experience changes the neural processing of action language. *Proc. Natl Acad. Sci. USA* **105**, 13 269–13 273. (doi:10.1073/pnas.0803424105)

18 Grafton, S., Fagg, A. & Arbib, M. 1998 Dorsal premotor cortex and conditional movement selection: a PET functional mapping study. *J. Neurophysiol.* **79**, 1092–1097.

19 Rushworth, M., Johansen-Berg, H., Gobel, S. & Devlin, J. 2003 The left parietal and premotor cortices: motor attention and selection. *Neuroimage* **20**, 89–100. (doi:10.1016/j.neuroimage.2003.09.011)

20 O'Shea, J., Sebastian, C., Boorman, E., Johansen-Berg, H. & Rushworth, M. 2007 Functional specificity of human premotor–motor cortical interactions during action selection. *Eur. J. Neurosci.* **26**, 2085–2095. (doi:10.1111/j.1460-9568.2007.05795.x)

21 Schluter, N., Krams, M., Rushworth, M. & Passingham, R. 2001 Cerebral dominance for action in the human brain: the selection of actions. *Neuropsychologia* **39**, 105–113. (doi:10.1016/S0028-3932(00)00105-6)

22 Toni, I., Shah, N. J., Frink, G. R., Thoenissen, D., Passingham, R. E. & Zilles, K. 2002 Multiple movement representations in the human brain: an event-related fMRI study. *J. Cogn. Neurosci.* **14**, 769–784. (doi:10.1162/08989290260138663)

23 Beattie, G. & Shovelton, H. 1999 Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *J. Lang. Soc. Psychol.* **18**, 438–462. (doi:10.1177/0261927X99018004005)

24 Kendon, A. 1994 Do gestures communicate? A review. *Res. Lang. Soc. Interact.* **27**, 175–200. (doi:10.1207/s15327973rlsi2703_2)

25 McNeill, D. 1992 *Hand and mind: what gestures reveal about thought.* Chicago, IL: University of Chicago Press.

26 Goldin-Meadow, S. & Sandhofer, C. M. 1999 Gesture conveys substantive information about a child's thoughts to ordinary listeners. *Dev. Sci.* **2**, 67–74. (doi:10.1111/1467-7687.00056)

27 Goldin-Meadow, S. & Singer, M. A. 2003 From children's hands to adults' ears: gesture's role in teaching and learning. *Dev. Psychol.* **39**, 509–520. (doi:10.1037/0012-1649.39.3.509)

28 Krauss, R. M., Morrel Samuels, P. & Colasante, C. 1991 Do conversational hand gestures communicate? *J. Pers. Soc. Psychol.* **61**, 743–754. (doi:10.1037/0022-3514.61.5.743)

29 Krauss, R. M., Dushay, R. A., Chen, Y. & Rausher, F. 1995 The communicative value of conversational hand gestures. *J. Exp. Soc. Psychol.* **31**, 533–552. (doi:10.1006/jesp.1995.1024)

30 Iverson, J. M. & Goldin-Meadow, S. 1998 Why people gesture when they speak. *Nature* **396**, 228. (doi:10.1038/24300)

31 Graham, J. A. & Heywood, S. 1976 The effects of elimination of hand gesture and of verbal codability on speech performance. *Eur. J. Soc. Psychol.* **5**, 189–195. (doi:10.1002/ejsp.2420050204)

32 Rauscher, F. B., Krauss, R. M. & Chen, Y. 1996 Gesture, speech and lexical access: the role of lexical movements in speech production. *Psychol. Sci.* **7**, 226–230. (doi:10.1111/j.1467-9280.1996.tb00364.x)

33 Goldin-Meadow, S., Nusbaum, H., Kelly, S. & Wagner, S. 2001 Explaining math: gesturing lightens the load. *Psychol. Sci.* **12**, 516–522. (doi:10.1111/1467-9280.00395)

34 Ping, R. & Goldin-Meadow, S. 2010 Gesturing saves cognitive resources when talking about non-present objects. *Cogn. Sci.* **34**, 602–619.

35 Wagner, S., Nusbaum, H. & Goldin-Meadow, S. 2004 Probing the mental representation of gesture: is hand-waving spatial? *J. Mem. Lang.* **50**, 395–407. (doi:10.1016/j.jml.2004.01.002)

36 Schwartz, D. L. & Black, J. B. 1996 Shuttling between depictive models and abstract rules: induction and fallback. *Cogn. Sci.* **20**, 457–498. (doi:10.1207/s15516709cog2004_1)

37 Goldin-Meadow, S. 2003 *Hearing gesture: how our hands help us think.* Cambridge, MA: The Belknap Press of Harvard University Press.

38 Perry, M., Church, R. B. & Goldin-Meadow, S. 1988 Transitional knowledge in the acquisition of concepts. *Cogn. Dev.* **3**, 359–400. (doi:10.1016/0885-2014(88)90021-4)

39 Kita, S. 2000 How representational gestures help speaking. In *Language and gesture* (ed. D. McNeill), pp. 162–185. New York, NY: Cambridge University Press.

40 Hostetter, A. B. & Alibali, M. W. 2008 Visible embodiment: gestures as simulated action. *Psychon. Bull. Rev.* **15**, 495–514. (doi:10.3758/PBR.15.3.495)

41 Streeck, J. 2002 A body and its gestures. *Gesture* **2**, 19–44. (doi:10.1075/gest.2.1.03str)

42 Cook, S. W. & Tanenhaus, M. K. 2009 Embodied communication: speakers' gestures affect listeners' actions. *Cognition* **113**, 98–104. (doi:10.1016/j.cognition.2009.06.006)

43 Newell, A. & Simon, H. A. 1972 *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

44 O'Reilly, A. W. 1995 Using representations: comprehension and production of actions with imagined objects. *Child Dev.* **66**, 999–1010. (doi:10.2307/1131794)

45 Boyatzis, C. J. & Watson, M. W. 1993 Preschool children's symbolic representation of objects through gestures. *Child Dev.* **64**, 729–735. (doi:10.2307/1131214)

46 Overton, W. F. & Jackson, J. P. 1973 The representation of imagined objects in action sequences: a developmental study. *Child Dev.* **44**, 309–314. (doi:10.2307/1128052)

47 Goodglass, H. & Kaplan, E. 1963 Disturbance of gesture and pantomime in aphasia. *Brain* **86**, 703–720. (doi:10.1093/brain/86.4.703)

48 Martin, P., Tewesmeier, M., Albers, M., Schmid, G. & Scharfetter, C. 1994 Investigation of gestural and pantomime performance in chronic schizophrenic inpatients. *Eur. Arch. Psychiatr. Clin. Neurosci.* **244**, 59–64. (doi:10.1007/BF02193520)

49 Frey, J. 2008 Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Phil. Trans. R. Soc. B* **363**, 1951–1957. (doi:10.1098/rstb.2008.0008)

50 Dumont, C., Ska, B. & Schiavetto, A. 1999 Selective impairment of transitive gestures: an unusual case of apraxia. *Neurocase* **5**, 447–458. (doi:10.1080/13554799908402739)

51 Rapcsak, S. Z., Ochipa, C., Beeson, P. M. & Rubens, A. B. 1993 Praxis and the right hemisphere. *Brain Cogn.* **23**, 181–202. (doi:10.1006/brcg.1993.1054)

52 Demir, Ö. E. 2009 A tale of two hands: development of narrative structure in children's speech and gesture and its relation to later reading skill. PhD dissertation, University of Chicago. (Proquest 3369323).

53 Broaders, S., Cook, S. W., Mitchell, Z. & Goldin-Meadow, S. 2007 Making children gesture brings out implicit knowledge and leads to learning. *J. Exp. Psychol. Gene.* **136**, 539–550. (doi:10.1037/0096-3445.136.4.539)

54 Ehrlich, S., Tran, K., Levine, S. & Goldin-Meadow, S. 2009 Gesture training leads to improvement in children's mental rotation skill. *Paper presented at the Biennial Meeting of the Society for Research in Child Development*, Denver.

55 Goldin-Meadow, S., Cook, S. W. & Mitchell, Z. A. 2009 Gesturing gives children new ideas about math. *Psychol. Sci.* **20**, 267–272. (doi:10.1111/j.1467-9280.2009.02297.x)

56 Beilock, S. & Goldin-Meadow, S. 2010 Gesture changes thought by grounding it in action. *Psychol. Sci.* **21**, 1605–1610. (doi:10.1177/0956797610385353)

57 Goldin-Meadow, S. & Beilock, S. L. 2010 Action's influence on thought: the case of gesture. *Perspect. Psychol. Sci.* **5**, 664–674. (doi:10.1177/1745691610388764)

58 Cartmill, E. A. & Maestripieri, D. 2011 Socio-cognitive specializations in nonhuman primates: evidence from gestural communication. In *The Oxford handbook of comparative evolutionary psychology* (eds J. Vonk & T. K. Shackelford), Oxford, UK: Oxford University Press.

59 Whiten, A., McGuigan, N., Marshall-Pescini, S. & Hopper, L. M. 2009 Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Phil. Trans. R. Soc. B* **364**, 2417–2428. (doi:10.1098/rstb.2009.0069)

60 Tomasello, M. 1996 Do apes ape? In *Social learning in animals: the roots of culture* (eds C. M. Heyes & B. G. Galef Jr), pp. 319–346. London, UK: Academic Press.

61 Whiten, A., Horner, V., Litchfield, C. A. & Marshall-Pescini, S. 2004 How do apes ape? *Learn. Behav.* **32**, 36–52. (doi:10.3758/BF03196005)

62 Nielsen, M., Collier-Baker, E., Davis, J. M. & Suddendorf, T. 2005 Imitation recognition in a captive chimpanzee (*Pan troglodytes*). *Anim. Cogn.* **8**, 31–36. (doi:10.1007/s10071-004-0232-0)

63 Dunbar, R. I. M., McAdam, M. R. & O'Connell, S. 2005 Mental rehearsal in great apes (*Pan troglodytes* and *Pongo pygmaeus*) and children. *Behav. Process.* **69**, 323–330. (doi:10.1016/j.beproc.2005.01.009)

64 Dunbar, R. I. M. 2000 Causal reasoning, mental rehearsal and the evolution of primate cognition. In *The evolution of cognition* (eds C. M. Heyes & L. Huber), pp. 205–219. Cambridge, MA: MIT Press.

65 Köhler, W. 1927 *The mentality of apes.* London, UK: Routledge & Kegan Paul [Transl. from 2nd revised edn by Ella Winter].

66 Kendon, A. 1991 Some considerations for a theory of language origins. *Man* **26**, 602–619. (doi:10.2307/2803829)

67 Suddendorf, T. & Busby, J. 2003 Mental time travel in animals? *Trends Cogn. Sci.* **7**, 391–396. (doi:10.1016/S1364-6613(03)00187-6)

68 Fagen, R. 2002 Primate juveniles and primate play. In *Juvenile primates* (eds M. E. Pereira & L. A. Fairbanks), Chicago, IL: University of Chicago Press.

69 Konner, M. 2010 *The evolution of childhood.* Cambridge, MA: The Belknap Press of Harvard University Press.

70 Cartmill, E. A. & Byrne, R. W. 2010 Semantics of primate gesture: determining intentional meanings. *Anim. Cogn.* **13**, 793–804. (doi:10.1007/s10071-010-0328-7)

71 Call, J. & Tomasello, M. (eds) 2007 *The gestural communication of apes and monkeys.* Mahwah, NJ: Lawrence Erlbaum.

72 Liebal, K., Pika, S. & Tomasello, M. 2004 Social communication in siamangs (*Symphalangus syndactulus*): use of gestures and facial expression. *Primates* **45**, 41–57. (doi:10.1007/s10329-003-0063-7)

73 Liebal, K., Pika, S. & Tomasello, M. 2006 Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture* **6**, 1–38. (doi:10.1075/gest.6.1.02lie)

74 Pika, S., Liebal, K. & Tomasello, M. 2003 Gestural communication in young gorillas (*Gorilla gorilla*): gestural repertoire, learning and use. *Am. J. Primatol.* **60**, 95–111. (doi:10.1002/ajp.10097)

75 Pika, S., Liebal, K. & Tomasello, M. 2005 Gestural communication in subadult bonobos (*Pan paniscus*): repertoire and use. *Am. J. Primatol.* **65**, 39–61. (doi:10.1002/ajp.20096)

76 Genty, E., Breuer, T., Hobaiter, C. & Byrne, R. W. 2009 Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim. Cogn.* **12**, 527–546. (doi:10.1007/s10071-009-0213-4)

77 Cartmill, E. A. & Byrne, R. W. 2011 Addressing the problems of intentionality and granularity in non-human primate gesture. In *Integrating gestures: the interdisciplinary nature of gesture* (eds G. Stam & M. Ishino), Amsterdam, The Netherlands: John Benjamins.

78 Liebal, K., Pika, S., Call, J. & Tomasello, M. 2004 Great ape communicators move in front of recipients

before producing visual gestures. *Interact. Stud.* **5**, 199–219. (doi:10.1075/is.5.2.03lie)

79 Poss, S., Kuhar, C., Stoinski, T. & Hopkins, W. D. 2006 Differential use of attentional and visual communicative signaling by orangutans (*Pongo pygmaeus*) and gorillas (*Gorilla gorilla*) in response to the attentional status of a human. *Am. J. Primatol.* **68**, 978–992. (doi:10.1002/ajp.20304)

80 Tanner, J. E. & Byrne, R. W. 1993 Concealing facial evidence of mood: perspective-taking in a captive gorilla. *Primates* **34**, 451–457. (doi:10.1007/BF02382654)

81 Leavens, D. A., Russell, J. L. & Hopkins, W. D. 2010 Multimodal communication by captive chimpanzees (*Pan troglodytes*). *Anim. Cogn.* **13**, 33–40. (doi:10.1007/s10071-009-0242-z)

82 Cartmill, E. A. & Byrne, R. W. 2007 Orangutans modify their gestural signaling according to their audience's comprehension. *Curr. Biol.* **17**, 1345–1348. (doi:10.1016/j.cub.2007.06.069)

83 Genty, E. & Byrne, R. 2010 Why do gorillas make sequences of gestures? *Anim. Cogn.* **13**, 287–301. (doi:10.1007/s10071-009-0266-4)

84 Liebal, K., Call, J. & Tomasello, M. 2004 Chimpanzee gesture sequences. *Primates* **64**, 377–396.

85 Klima, E. & Bellugi, U. 1979 *The signs of language.* Cambridge, MA: Harvard University Press.

86 Kendon, A. 2004 *Gesture: visible action as utterance.* Cambridge, UK: Cambridge University Press.

87 Feldman, H., Goldin-Meadow, S. & Gleitman, L. 1978 Beyond Herodotus: the creation of language by linguistically deprived deaf children. In *Action, symbol, and gesture: the emergence of language* (ed. A. Look), New York, NY: Academic Press.

88 Goldin-Meadow, S. 2003 *The resilience of language: what gesture creation in deaf children can tell us about how all children learn language.* New York, NY: Psychology Press.

89 Goldin-Meadow, S. & Feldman, H. 1977 The development of language-like communication without a language model. *Science* **197**, 401–403. (doi:10.1126/science.877567)

90 Goldin-Meadow, S. & Mylander, C. 1983 Gestural communication in deaf children: the non-effects of parental input on language development. *Science* **221**, 372–374. (doi:10.1126/science.6867713)

91 Goldin-Meadow, S. & Mylander, C. 1998 Spontaneous sign systems created by deaf children in two cultures. *Nature* **91**, 279–281. (doi:10.1038/34646)

92 Gershkoff-Stowe, L. & Goldin-Meadow, S. 2002 Is there a natural order for expressing semantic relations? *Cogn. Psychol.* **45**, 375–412. (doi:10.1016/S0010-0285(02)00502-9)

93 Goldin-Meadow, S., McNeill, D. & Singleton, J. 1996 Silence is liberating: removing the handcuffs on grammatical expression in the manual modality. *Psychol. Rev.* **103**, 34–55. (doi:10.1037/0033-295X.103.1.34)

94 Goldin-Meadow, S., So, W.-C., Ozyurek, A. & Mylander, C. 2008 The natural order of events: how speakers of different languages represent events nonverbally. *Proc. Natl Acad. Sci. USA* **105**, 9163–9168. (doi:10.1073/pnas.0710060105)

95 Hall, M., Mayberry, R. & Ferreira, V. 2010 Communication systems shape the natural order of events: competing biases from grammar and pantomime. In *Abstracts of the 4th Conference of the International Society for Gesture Studies*, p. 182. Amsterdam, The Netherlands: John Benjamins.

96 Langus, A. & Nespor, M. 2010 Cognitive systems struggling for word order. *Cogn. Psychol.* **60**, 291–318. (doi:10.1016/j.cogpsych.2010.01.004)

97 Meir, I., Lifshitz, A., Ilkbasaran, D. & Padden, C. 2010 The interaction of animacy and word order in human languages: a study of strategies in a novel communication task. In *Proc. 8th Evolution of Language Conf.* (eds A. D. M. Smith, M. Schouwstra, B. de Boer & K. Smith), pp. 455–456. Singapore, Singapore: World Scientific.

98 Tomasello, M. & Camaioni, L. 1997 A comparison of the gestural communication of apes and human infants. *Hum. Dev.* **40**, 7–24. (doi:10.1159/000278540)

99 Leavens, D. A. 2004 Manual deixis in apes and humans. *Interact. Stud.* **5**, 387–408. (doi:10.1075/is.5.3.05lea)

100 Leavens, D. A., Hopkins, W. D. & Bard, K. 1996 Indexical and referential pointing in chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* **110**, 346–353. (doi:10.1037/0735-7036.110.4.346)

101 Pika, S. & Mitani, J. C. 2006 Referential gesturing in wild chimpanzees (*Pan troglodytes*). *Curr. Biol.* **16**, 191–192. (doi:10.1016/j.cub.2006.02.037)

102 Liebal, K. & Call, J. 2012 The origins of nonhuman primates' manual gestures. *Phil. Trans. R. Soc. B* **367**, 118–128. (doi:10.1098/rstb.2011.0044).

103 Savage-Rumbaugh, E. S., Wilkerson, B. J. & Bakeman, R. 1977 Spontaneous gestural communication among conspecifics in the pygmy chimpanzee (*Pan paniscus*). In *Progress in ape research* (ed. G. H. Bourne), pp. 97–116. New York, NY: Academic Press.

104 Tanner, J. E. & Byrne, R. W. 1996 Representation of action through iconic gesture in a captive lowland gorilla. *Curr. Anthropol.* **37**, 162–173. (doi:10.1086/204484)

105 Tanner, J. E. & Byrne, R. W. 1999 The development of spontaneous gestural communication in a group of zoo-living lowland gorillas. In *The mentalities of gorillas and orangutans: comparative perspectives* (eds S. T. Parker, R. W. Mitchell & H. L. Miles), pp. 211–239. Cambridge, UK: Cambridge University Press.

106 Seyfarth, R. M. & Cheney, D. L. 2003 Signalers and receivers in animal communication. *Annu. Rev. Psychol.* **54**, 145–173. (doi:10.1146/annurev.psych.54.101601.145121)

107 Rizzolatti, G. & Craighero, L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)

108 Hopper, L. M., Spiteri, A., Lambeth, S. P., Schapiro, S. J., Horner, V. & Whiten, A. 2007 Experimental studies of traditions and underlying transmission processes in chimpanzees. *Anim. Behav.* **73**, 1021–1032. (doi:10.1016/j.anbehav.2006.07.016)

109 van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C. D., Singleton, I., Suzuki, A., Utami, S. S. & Merrill, M. 2003 Orangutan cultures and the evolution of material culture. *Science* **299**, 102–105. (doi:10.1126/science.1078004)

110 Whiten, A., Goodall, J., McGrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., Tutin, C. E. G., Wrangham, R. W. & Boesch, C. 1999 Cultures in chimpanzees. *Nature* **399**, 682–685. (doi:10.1038/21415)

111 Gardner, R. A. & Gardner, B. T. 1969 Teaching sign language to a chimpanzee. *Science* **165**, 664–672. (doi:10.1126/science.165.3894.664)

112 Miles, H. L. W. 1990 The cognitive foundations for reference in a signing orangutan. In *'Language' and intelligence in monkeys and apes. Comparative developmental perspectives* (eds S. T. Parker & K. R. Gibson), pp. 511–539. Cambridge, UK: Cambridge University Press.

113 Patterson, F. G. & Linden, E. 1981 *The education of Koko.* New York, NY: Owl Books.

114 Savage-Rumbaugh, E. S., Shanker, S. G. & Taylor, T. J. 1998 *Apes, language, and the human mind.* New York, NY: Oxford University Press.

115 Fouts, R. & Mills, S. 1997 *Next of kin: what chimpanzees have taught me about who we are.* New York, NY: Morrow.

116 Greenfield, P. M. & Savage-Rumbaugh, E. S. 1991 Imitation, grammatical development, and the invention of protogrammar by an ape. In *Biological and behavioral determinants of language development* (eds N. A. Krasnegor, D. M. Rumbaugh, R. L. Schiefelbusch & M. Studdert-Kennedy), pp. 235–262. Hillsdale, NJ: Earlbaum Associates.

117 Premack, A. J. & Premack, D. 1972 Teaching language to an ape. *Sci. Am.* **227**, 92–99. (doi:10.1038/scientificamerican1072-92)

118 Premack, A. J. & Premack, D. 1983 *The mind of an ape.* New York, NY: Norton.

119 Snowdon, C. T. & Hausberger, M. 1997 *Social influences on vocal development.* Cambridge, UK: Cambridge University Press.

120 Armstrong, D. F., Stokoe, W. C. & Wilcox, S. E. 1995 *Gesture and the nature of language.* Cambridge, UK: Cambridge University Press.

121 Armstrong, D. F. & Wilcox, S. E. 2007 *The gestural origin of language.* Oxford, UK: Oxford University Press.

122 Hewes, G. W. 1973 Primate communication and the gestural origin of language. *Curr. Anthropol.* **12**, 5–24. (doi:10.1086/201401)

123 Corballis, M. C. 2002 *From hand to mouth, the origins of language.* Princeton, NJ: Princeton University Press.

124 Iverson, J. M. & Goldin-Meadow, S. 2005 Gesture paves the way for language development. *Psychol. Sci.* **16**, 367–371. (doi:10.1111/j.0956-7976.2005.01542.x)

125 Ozcaliskan, S. & Goldin-Meadow, S. 2005 Gesture is at the cutting edge of early language development. *Cognition* **96**, 101–113. (doi:10.1016/j.cognition.2005.01.001)

126 Arbib, M. A., Liebal, K. & Pika, S. 2008 Primate vocalization, gesture, and the evolution of human language. *Curr. Anthropol.* **49**, 1053–1076. (doi:10.1086/593015)

127 Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. 1992 Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180.

128 Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. 1996 Action recognition in the premotor cortex. *Brain* **119**, 593–609. (doi:10.1093/brain/119.2.593)

129 Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. 1996 Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* **3**, 131–141. (doi:10.1016/0926-6410(95)00038-0)

130 Rizzolatti, G. & Sinigaglia, C. 2010 The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nat. Rev. Neurosci.* **11**, 264–274. (doi:10.1038/nrn2805)

131 Rizzolatti, G. & Luppino, G. 2001 The cortical motor system. *Neuron* **31**, 889–901. (doi:10.1016/S0896-6273(01)00423-8)

132 Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. & Rizzolatti, G. 2001 'I know what you are doing': a neurophysiological study. *Neuron* **32**, 91–101.

133 Kraskov, A., Dancause, N., Quallo, M. M., Shepherd, S. & Lemon, R. N. 2009 Corticospinal neurons in macaque ventral premotor cortex with mirror properties: a potential mechanism for action suppression? *Neuron* **64**, 922–930. (doi:10.1016/j.neuron.2009.12.010).

134 Ferrari, P. F., Gallese, V., Rizzolatti, G. & Fogassi, L. 2003 Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* **17**, 1703–1714. (doi:10.1046/j.1460-9568.2003.02601.x)

135 Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V. & Rizzolatti, G. 2002 Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846–848. (doi:10.1126/science.1070311)

136 Buccino, G. *et al.* 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404.

137 Grezes, J., Armony, J. L., Rowe, J. & Passingham, R. E. 2003 Activations related to 'mirror' and 'canonical' neurons in the human brain: an fMRI study. *Neuroimage* **18**, 928–937. (doi:10.1016/S1053-8119(03)00042-9)

138 Skipper, J. A., Goldin-Meadow, S., Nusbaum, H. C. & Small, S. L. 2007 Speech-associated gestures, Broca's area, and the human mirror system. *Brain Lang.* **101**, 260–277. (doi:10.1016/j.bandl.2007.02.008)

139 Arbib, M. A. 2005 From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* **28**, 105–167.

140 Leavens, D. A., Hopkins, W. D. & Bard, K. 2005 Understanding the point of chimpanzee pointing: epigenesis and ecological validity. *Curr. Direct. Psychol. Sci.* **14**, 185–189. (doi:10.1111/j.0963-7214.2005.00361.x)

141 Lyn, H., Russell, J. L. & Hopkins, W. D. 2009 The impact of environment on the comprehension of declarative communication in apes. *Psychol. Sci.* **21**, 360–365. (doi:10.1177/0956797610362218)

142 Harvey, P. H. & Clutton-Brock, T. H. 1985 Life history variation in primates. *Evolution* **39**, 559–581. (doi:10.2307/2408653)

143 Rowe, M. L. & Goldin-Meadow, S. 2009 Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science* **323**, 951–953. (doi:10.1126/science.1167025)

144 Goldin-Meadow, S., Goodrich, W., Sauer, E. & Iverson, J. 2007 Young children use their hands to tell their mothers what to say. *Dev. Sci.* **10**, 778–785. (doi:10.1111/j.1467-7687.2007.00636.x)

Review

A neuropsychological perspective on the link between language and praxis in modern humans

Agnes Roby-Brami[1,2,*], **Joachim Hermsdörfer**[3], **Alice C. Roy**[4,6] **and Stéphane Jacobs**[5,6]

[1]*Laboratory of Neurophysics and Physiology, University Paris Descartes, CNRS UMR 8119, 45 rue des Saints Pères, 75006 Paris, France*
[2]*Department of Physical Medicine and Rehabilitation, AP-HP, Raymond Poincaré Hospital, 92380 Garches, France*
[3]*Department of Sport and Health Science, Technische Universität München, Georg-Brauchle-Ring 60-62, 80992 München, Germany*
[4]*CNRS FRE 3406, Laboratory on Language, Brain and Cognition, Institute for Cognitive Sciences, 67 Boulevard Pinel, 69675 Bron, France*
[5]*INSERM U1028, CNRS UMR5292, Lyon Neuroscience Research Center, ImpAct team, 16 Avenue Lépine, 69676 Bron, France*
[6]*University Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne, France*

Hypotheses about the emergence of human cognitive abilities postulate strong evolutionary links between language and praxis, including the possibility that language was originally gestural. The present review considers functional and neuroanatomical links between language and praxis in brain-damaged patients with aphasia and/or apraxia. The neural systems supporting these functions are predominantly located in the left hemisphere. There are many parallels between action and language for recognition, imitation and gestural communication suggesting that they rely partially on large, common networks, differentially recruited depending on the nature of the task. However, this relationship is not unequivocal and the production and understanding of gestural communication are dependent on the context in apraxic patients and remains to be clarified in aphasic patients. The phonological, semantic and syntactic levels of language seem to share some common cognitive resources with the praxic system. In conclusion, neuropsychological observations do not allow support or rejection of the hypothesis that gestural communication may have constituted an evolutionary link between tool use and language. Rather they suggest that the complexity of human behaviour is based on large interconnected networks and on the evolution of specific properties within strategic areas of the left cerebral hemisphere.

Keywords: action; language; brain damage; gesture; pantomime; tool use

1. INTRODUCTION

Language and complex actions or praxis, including tool use, are cognitive functions that, although present to some degree in many animal species, are uniquely developed in humans. In addition to being distinctive human traits, these two behaviours are mainly controlled by the left cerebral hemisphere in the vast majority of individuals, as demonstrated by neuropsychological observations. This lateralization is reminiscent of the very strong population-level bias for dextrality in the human species, whereby approximately 90 per cent of individuals favour their right hand for fine motor skills [1]. These converging cerebral asymmetries have led researchers to consider the left hemisphere as dominant for language as well as for motor functions [2], and have triggered interest in the potential evolutionary and functional links between manual preference, tool use and language.

The origin of the left hemisphere specialization for language and praxis, including tool use, and its relation to manual preference, is still disputed. For example, some argue that dextrality might have emerged first [3], while others propose that it appeared under selective pressure for common handedness as an advantage for learning tool use through imitation [4]. Other authors suggest that human dextrality is simply a mere consequence of the ancient left lateralization of the cerebral control of vocalization, as seen in many species from birds to mammals. According to this hypothesis, the progressive incorporation of vocalization into an

* Author for correspondence (agnes.roby-brami@parisdescartes.fr).

originally gestural language would have led to a left hemispheric specialization for language and motor control [5].

Regardless of its relationship with manual preference, the study of the link between language and tool use is highly relevant to the understanding of the development of these two unique human abilities and the origins of our species. According to archaeological records, tool use emerged about 2.5 Ma, starting with simple behaviours such as modifying rocks for pounding [6] and then progressing towards the construction of more and more refined and complex compound tools through cumulative evolution [7]. In parallel, language is thought to have emerged owing to the social interactions required by the development of human technology, in particular by learning tool-related behaviours through imitation [8]. An increasingly hierarchical organization of language would then have appeared thanks to a pre-existing left hemispheric specialization for hierarchically and sequentially ordered behaviours, initially developed for the manufacture and use of tools [4]. Developmental studies investigating language and object combination behaviours in young children, as well as work carried out in primates and apes, suggest that language and tool use do indeed share some common functional and neural foundations both phylogenetically and ontogenetically during the first years of development [9].

The cerebral basis of tool use in monkeys as well as in humans has been extensively investigated over the past two decades. Iriki *et al.* [10] first demonstrated that simple tool use, i.e. using a rake to retrieve food placed out of reach, is accompanied in macaques by plastic changes of sensory responses of neurons in the parietal cortex. This seminal work, together with subsequent studies done in monkeys (e.g. [11]), led Frey [12] to propose that simple tool-use behaviours, in which the tool merely constitutes a functional extension of the limb [13,14], rely on experience-dependent changes in areas within the dorsal stream of visual processing [15,16], known to be essentially involved in sensory-motor transformations for the control of actions [17]. Recent studies in humans [18] and in monkeys [19] support this hypothesis. In contrast with simple tool use, complex tool use, like most everyday familiar actions, is a uniquely human skill whereby the use of a tool 'converts the movements of the hands into qualitatively different mechanical actions' [20]. This ability depends not only on sensory-motor transformations for the control of action, but also on access to acquired semantic knowledge about the tool and its common uses [21]. So, complex tool use draws upon the collaboration between the aforementioned dorsal stream and the ventral visual pathway [12] thought to be responsible for object recognition and the building and storage of semantic knowledge [16]. Accordingly, data obtained from brain imaging studies of various complex tool-use tasks in able-bodied subjects show that these behaviours recruit a large distributed network within the temporal, parietal and frontal areas, primarily lateralized to the left hemisphere [22,23]. Further evidence for this integration of semantics into sensory-motor control of action can be found in the fact that conceptual knowledge influences the way people

spontaneously grasp familiar tools [24]. Importantly, this effect can be disrupted in patients with left but not right side brain damage [25]. Based on these findings, it has been proposed that the unique human abilities of designing and using complex tools originate from adaptations of sensory-motor networks and their integration with cognitive processes pertaining to semantic knowledge about tools, the agent's intentions and contextual information about the task, most of these being also supported by the left hemisphere [26].

The emergence of language, on the other hand, is often conceived of as depending critically on the receiver's ability to decode the sender's message or intentions, subserved by some common representations between the two [27]. The discovery of mirror neurons in the monkey [28] might have provided the link between action execution and recognition that is necessary for communication in general. Mirror neurons fire not only when the monkey executes specific grasping actions, but also when it perceives the same action being performed by another individual. These neurons have been observed in area F5 of the ventral premotor cortex of macaques as well as in the inferior parietal lobule, where some of these neurons also show sensitivity to the goal of the action, independently from the motor details of its execution [29]. These two brain regions are known to be reciprocally connected, and are part of the dorsal visual stream subserving the sensory transformations involved in the control of reaching and grasping actions. Interestingly, in the context of the emergence of language, the putative human homologue of area F5 is the caudal part of the inferior frontal gyrus, which corresponds, on the left side, to Broca's area, known for its involvement in many aspects of language, from phonology to syntax and from production to comprehension [30,31]. In addition, evidence for the existence of a mirror system in humans has been reported [32], providing a possible neural basis for action understanding [33]. This human analogue of the monkey mirror neuron system may in addition support a variety of complex socio-cognitive phenomena, including language [34], although this view is challenged by recent work [35,36]. Regarding the evolution of language, and following Liberman's proposal mentioned earlier [27], the mirror neuron system would thus have allowed the mapping of the sender's message and intentions onto the receiver's own representations, laying the bases for a primitive gestural form of language [5,37]. If this is the case, then Broca's area as we know it now would have developed 'atop a mirror neuron system for grasping' through increasingly complex stages of gesture recognition and imitation [38].

Independently from the potential involvement of a mirror neuron system, the relationship between praxis, gesture and language has to be further examined on the basis of recent neuropsychological data [39]. The lateralization to the left hemisphere seems to be the key phenomenon for evolution of both language and complex action systems in humans. Indeed, clinical observations gathered for more than a century have demonstrated that a lesion of the left hemisphere may induce a disruption of language (aphasia) [40] and of complex action systems (apraxia) [41]. These disorders

are very often associated in brain-damaged patients [42]. Classical neuropsychological analyses rely on the clinical dissociation of the elementary impairments constituting the aphasic and apraxic syndromes and their confrontation with post-mortem neuroanatomy in order to describe brain–behaviour relationships. This classical approach has led to the elaboration of cognitive models of language and action. Nowadays, neuropsychology has largely benefited from progress in brain imaging, which allows precise investigations of the neural bases of higher brain functions and the mechanisms of their dysfunction.

Here, we will first outline the clinical picture and the first interpretations of aphasia and apraxia. In §3, we will present the more contemporary theoretical accounts of these disorders, contrasting the early localizationist approaches with current views that these behaviours are supported by widespread, dynamic neural networks. Then, we will examine the link between praxis and language by reviewing the effects of brain lesions on several relevant behaviours such as action recognition, repetition and imitation, and gestural communication. We will attempt to compare these alterations despite the fact that most clinical studies in the literature focus on either aphasia or apraxia, and use generally different clinical approaches and different theoretical backgrounds. Finally, we will examine the possibility that both action and language share common cognitive resources.

2. CLINICAL DESCRIPTION AND ORIGINAL ACCOUNTS OF APHASIA AND APRAXIA

(a) *Early interpretations of aphasia and the concept of brain localization*

Language refers to a system of signs (indices, icons, symbols) used to encode and decode information so that the pairing of a specific sign with an intended meaning is established through social conventions. Language presents several aspects: phonological, semantic, syntactic, prosodic and pragmatic, which can be differentially impaired after brain lesions [43]. The phonological level refers to the sounds used in the language. Each language thus has a different phonology, as certain sounds will be present in one language but not in another. Semantics refers to the meaning of language, and syntax represents the principles and rules for constructing sentences. The phonological, semantic and syntactic aspects of language are to a vast extent specific to humans. Prosody refers to the voice modulation that accompanies different emotional content or intention, and is classically attributed to the right hemisphere [44,45]. Finally, the pragmatic aspect of language refers to the complex combinations of symbols used to transmit complex ideas and includes many other cognitive functions, supported by both hemispheres [46].

Aphasia corresponds to impairment, following a brain lesion, of phonological, syntactic and/or semantic processing, either in isolation or in association, and may concern either language production or comprehension, or both. These three aspects of language usually being essentially supported by the left hemisphere in right-handers [47], aphasia follows left brain damage in the vast majority of patients [48].
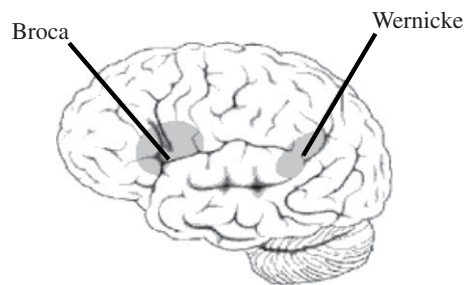
Figure 1. Localization of Broca's and Wernicke's areas.

This is an acquired disorder: the term excludes developmental language disorders in children.

(i) *The concept of localization*

Historically, the topic of aphasia was at the centre of the debate between localized versus holistic explanations of psychological functions of the brain. Franz Joseph Gall was the first to propose separate brain localizations for different behaviours. Broca [40] presented the first clinical case in which focal brain damage was associated with altered language production (figure 1). Later, Carl Wernicke described the defect of language comprehension after a lesion of the posterior section of the superior temporal gyrus. The localization of language functions was then challenged by the holistic theory, which posited a single language function performed by the left hemisphere (review in [49]). Geschwind [50,51] reconsidered localization and proposed that the impairments were the result of disconnection between brain areas (review in [52]).

(ii) *Types of aphasia*

Since the work of Broca & Wernicke in the nineteenth century, the definition and different types of aphasia [53] have been refined. Broca's aphasia (also referred to as non-fluent or agrammatic aphasia) is caused by damage to anterior regions of the brain, in particular to Broca's area, corresponding to the caudal part of the left inferior frontal gyrus (Brodmann areas 44 and 45). It is characterized by reduced, non-fluent agrammatical spontaneous speech with relatively spared comprehension. Fluency impairments include reduced phrase length, altered melody and articulation, reduced word flow or agrammatical sentences. Some over-learned social phrases may paradoxically be preserved and fluent. Comprehension is usually preserved as long as simple, semantically non-reversible sentences are used; however, patients' performance may drastically drop when tested with syntactically complex sentences [54–57]. The severity of Broca's aphasia varies greatly. When the vascular damage includes the anterior insula, the linguistic deficit is accompanied by a motor deficit (the so-called apraxia of speech) characterized by disrupted articulation and prosody [58–61]. Wernicke's aphasia (also called fluent aphasia), on the other hand, is caused by neurological damage to the posterior part of the superior temporal gyrus (Brodmann area 22). It is characterized by paragrammatic, fluent but relatively meaningless spontaneous speech expressed with the appropriate melody or intonation. Spoken language may be limited to jargon with many neologisms, paraphasias or non-words. The comprehension
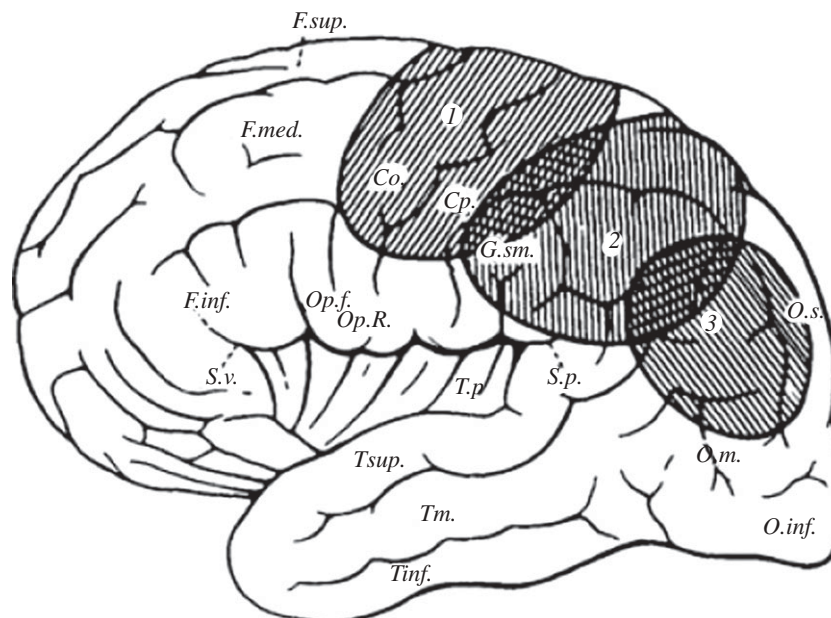
Figure 2. Topography of left hemisphere lesions leading to the three types of apraxia according to Liepmann [41]. 1: Melokinetic apraxia; 2: ideokinetic or ideomotor apraxia; 3: ideational apraxia.

of words, sentences and conversations is relatively poor: patients are typically not aware of their errors. A combination of Broca's and Wernicke's aphasia may be observed in the case of large lesions. The major language impairment observed in those patients is referred to as global aphasia.

Aphasia can also occur without damage to Broca's or Wernicke's areas. Transcortical aphasias are thus due to lesions surrounding these areas. The respective language syndromes are similar to Broca's or Wernicke's aphasia, except that word repetition is preserved. By contrast, conduction aphasia is characterized by a predominant impairment of word repetition, as well as frequent phonemic paraphasias with unsuccessful attempts at self-correction ('*conduite d'approche*'), naming difficulties with relatively well-preserved auditory comprehension and fluent, grammatically correct spontaneous speech production. This particular pattern of deficit, leaving unimpaired the linguistic comprehension and production, has led Ardila [62] to suggest that conduction aphasia might not be a linguistic deficit *per se*. This conception is not new. Indeed, Luria [63] proposed to interpret conduction aphasia as a segmental ideomotor apraxia. Along the same line, Benson *et al.* [64] reported ideomotor apraxia as a secondary characteristic of conduction aphasia. We will come back to the specific case of conduction aphasia and its relation to apraxia shortly. Regardless, while conduction aphasia has been traditionally viewed as a disconnection between Broca's and Wernicke's areas owing to damage of the arcuate fasciculus [52], recent brain imaging studies have underlined the role of the supramarginal gyrus and neighbouring cortical territories in word repetition [65,66]. This region of the brain is also regarded as central in the cerebral organization of praxis.

(b) Apraxia and the localization of higher motor functions

Apraxia is a disorder of learned, purposive skilled movement that is not explained by deficits of the elemental motor or sensory systems, or by general cognitive impairment [67]. The symptoms are bilateral although they are caused by unilateral, predominantly left-sided, brain lesions. The main symptoms of apraxia are most obvious during performance of meaningful gestures as recognized since Liepmann's original description [41]. The literature on the subject of apraxia distinguishes two types of meaningful gestures: 'transitive' gestures, which involve a tool or an object (including tool-use pantomimes), and 'intransitive' gestures, which, in fact, are mainly symbolic and communicative (e.g. waving good-bye). Impairments may vary according to the mode of elicitation of the action (i.e. executed either on command or on imitation), and may also affect action recognition. Specific impairments can be selectively observed in the case of focal lesions, but they are more frequently combined. The skilfulness of the patient's movements depends on the conditions and context of their elicitation [68].

(i) Types of apraxia according to Liepmann

Liepmann proposed that performing a gesture is based on the collaborative interaction of central processes. From a visuokinaesthetic image of an intended motor act, a 'formula' of movement is derived within the left posterior cortical areas. During gesture performance, this representation is 'transcoded' to activate the appropriate muscle groups supported by 'kinaesthesic memories' of learned movements stored in the sensorimotor cortex. This step requires the existence of intact connections between the posterior cortical areas and the sensorimotor cortex [41].

Liepmann distinguished three types of apraxia (figure 2). *Ideational apraxia* corresponds either to a disturbance in the 'movement formula' or to a lack of access to this representation. It is characterized by inadequate use of objects or by the wrong arrangement of the various steps of sequential actions. *Ideomotor apraxia* corresponds to a disconnection between the formula and 'kinaesthesic memories' owing to a lesion

within the left posterior parietal lobe. The 'innervatory patterns' are preserved, but their activation by the formula is impossible or impaired. It is characterized by adequate movements, performed in response to a command or self-generated, but their performance is degraded. *Melokinetic apraxia* corresponds to a generalized clumsiness owing to a lesion of kinaesthetic 'memories' stored in the sensorimotor cortex.

(ii) *Other neuropsychological models of apraxia*

Geschwind [50,51] reconsidered the question of localization and confirmed that in right-handed persons, the left hemisphere is dominant for complex gestures. However, he focused more on the importance of white matter lesions, than on stored representation of gestures. For him, the inability to pantomime the use of an object upon verbal command is rather the consequence of the disconnection of frontal premotor areas from Wernicke's area owing to a lesion of the left arcuate fasciculus. Geschwind [50] proposed the same interpretation and anatomical correlates to account for conduction aphasia, a syndrome that bears many functional and anatomical resemblances with ideomotor apraxia.

More recently and in line with Liepmann's original proposal, Roy & Square [69] proposed a two-system action model: a *conceptual* system, including semantic knowledge of tools, objects and actions, and a *production* system representing the sensorimotor knowledge of action as well as perceptuo-motor processes that allow its organization and implementation. The conceptual system defines the action plan according to the knowledge of objects and tools, the context-independent knowledge of action and the knowledge of the arrangement of simple actions in a sequence. The production system includes motor programmes independent of the effectors, which permit the action to be carried out according to the context and needs. Praxic disturbances can thus be interpreted in terms of impairment of the conceptual system (ideational apraxia) and/or the production system of action (ideomotor apraxia). In ideomotor apraxia, knowledge pertaining to objects and tools is preserved and patients can therefore describe and identify actions associated with tools and appreciate their adequacy, while being unable to perform them adequately.

Heilman & Rothi (review in [70,71]) proposed a cognitive model inspired by models of the language system, in order to account for all the dissociations observed in patients depending on which modality is used to elicit gestures (verbal command, presentation of objects, imitation, etc.). This model is constituted of several modules, which process specific information and are centred on an action semantic system. They propose that sensory information accesses the system via an action input lexicon that contains information about the physical attributes of perceived actions (mainly visual representations). The semantic action system then integrates information transferred from the action input lexicon and is at least partially independent from other forms of semantic knowledge. The action output lexicon subsequently includes information pertaining to the physical attributes of an action to be performed (mainly kinaesthetic representations). Apart from this indirect lexical route, a direct, non-lexical route, based

mainly on visual processing of perceived gestures, controls the imitation of meaningless or unfamiliar gestures, with a possible dissociation between those routes. Based on the observation that visual recognition of action and movement can be impaired in some apraxic patients with posterior lesions, Heilman *et al.* [72] proposed that two forms of ideomotor apraxia exist: one owing to posterior lesions, destroying the areas containing visuokinaesthetic engrams (and thus also impairing gesture recognition) and the other owing to more frontal lesions, potentially disconnecting motor areas from visuokinaesthetic engrams, therefore preserving gesture recognition.

De Renzi & Luchelli [73] investigated ideational apraxia with specific tests: multiple-step tasks, and tasks requiring the understanding of tool and/or object properties (tool selection, alternative tool selection, gesture recognition). They found that the scores for multiple and single tool-use tasks were correlated with each other but not with the results of a test assessing ideomotor apraxia. They concluded that 'ideational apraxia is an autonomous syndrome, linked to left hemisphere damage and pertaining to the area of semantic memory disorders rather than to that of defective motor control'.

An important characteristic of apraxia is the well-known 'automatic/voluntary' dissociation whereby patients fail to perform adequate gestures on command while their performance on similar self-initiated actions in daily life is preserved, showing that the full context of action is particularly important for the retrieval and execution of adequate gestures. This dissociation has been confirmed by experimental methods [74]. This observation is reminiscent of the relative sparing of over-learned social phrases described in non-fluent aphasic patients [53].

Early theories of the cerebral bases of praxis and language have thus focused on localizing different aspects of these faculties to specific areas, interpreting apraxic and aphasic disorders in terms of either damage to one of these areas or disconnection between them. However, these conceptions have failed to account for many disorders exhibited by brain-damaged patients, leading to the emergence of more complex and integrated conceptions of the brain bases of these complex cognitive faculties.

3. CONTEMPORARY THEORIES FOR THE CEREBRAL ORGANIZATION OF LANGUAGE AND PRAXIS

Beyond the neuropsychological approach, more recent work has attempted to explain the different manifestations of apraxia as dysfunctions of the sensorimotor systems for action and object manipulation, as recently identified by electrophysiological studies in monkeys [75]. In particular, Buxbaum [76] proposed a model of gesture production aimed at reconciling the classical neuropsychological interpretations of apraxic disorders with this more recent neurophysiological framework (review in [77]). In line with classical neuropsychological models [41,78], this model assumes the existence of gesture engrams, which are conceived of as memory-stored sensorimotor (i.e. non-verbal) representations of familiar gestures, involved in both gesture production and

recognition [72]. These engrams are thought to be stored in the left inferior parietal lobule [79] at the interface between the ventral and dorsal streams of visual processing. The evocation of a familiar gesture, for example by verbal command or presentation of the associated tool, would thus activate the appropriate gestural engrams via the lexical semantic system, located primarily in ventral regions. Gesture execution would then be controlled by a dynamic system constituted of the parieto-frontal networks of the dorsal stream. In accordance with this view, patients with damage to the inferior parietal lobule may be able to normally grasp objects based solely on their physical properties (i.e. affordances), while being impaired when asked to grasp the same objects in order to use them, which requires integration of conceptual knowledge of tools and their function [80,81]. More recently, Frey [26] proposed a more dynamic conception of the role of the left inferior parietal cortex. Rather than storing gesture engrams, this region would assemble praxis representations in order to fit all the constraints imposed by conceptual knowledge about tools and their function, the task context, the agent's intentions, etc.

Recent conceptions on the functional anatomy of language have also largely departed from the classical localizationist views exposed in §2, now favouring the idea that language might be organized in networks rather than specialized brain areas. Indeed, the classical concept according to which the frontal lobe is responsible for speech production and temporal areas for language comprehension could not account for the cases of patients presenting, for example, a syndrome of Broca's aphasia with no lesions to Broca's area, or deficits of speech comprehension associated with a lesion in Broca's region [54,55,57]. Furthermore, recent studies have confirmed, for example, the involvement of Broca's area in language comprehension in healthy individuals, at the phonological [31,82,83], lexical [31,84] and syntactical levels [85–87]. These observations, together with the development of neuroimaging techniques, have led researchers to consider that different, partially overlapping networks of superior temporal, posterior parietal and ventral prefrontal areas underpin the phonological, semantic and syntactic levels of language [47].

Current theoretical accounts of language and praxis thus favour the conception that these complex cognitive faculties are subserved by neural networks widely distributed in the left hemisphere. The recruitment of different neural systems would then depend on the exact nature, constraints and context of the task [26,88]. The multiple aspects of aphasia and apraxia would therefore result from disruption of, or the imbalance in, the interactions between parts of these networks, rather than from localized damage to a brain area supporting a specific function. Regardless of this evolution of the theoretical framework for the functional neuroanatomy subserving language and praxis and their disorders, it appears clearly that both cognitive abilities rely on largely overlapping networks, with critical nodes located in the superior temporal, rostral inferior parietal and ventral premotor cortices.

We now turn to examine the implications of language and praxic disorders for action recognition, imitation and gestural communication, which, as we highlighted in §1, might have constituted critical abilities for the propagation of human technology and language evolution.

4. ACTION RECOGNITION, IMITATION AND GESTURAL COMMUNICATION IN APHASIA AND APRAXIA

(a) *Action recognition*

As mentioned above, Heilman *et al.* [72] showed that the recognition of transitive tool-use gestures was impaired in some patients with ideomotor apraxia, who were therefore also impaired for the execution of these gestures. These findings were later confirmed [89,90]. Neuroimaging experiments also showed a similar interaction between action observation and production (and imagination) in healthy humans [91]. The observation of meaningful actions activates the left hemisphere in the frontal and temporal regions while the observation of meaningless actions involves mainly the right occipito-parietal pathway [92].

The link between observation and imitation of object-related actions in apraxic patients has been recently re-examined by Buxbaum *et al.* [93] and Pazzaglia *et al.* [94] with advanced lesion reconstruction techniques. Buxbaum *et al.* observed a close relationship between performance in pantomime recognition tasks and imitation of object-related actions. Further, in line with Heilman *et al.* [72], the neuroanatomical analysis showed that lesions located in the inferior parietal lobe and in the intraparietal sulcus were significantly associated with deficits in the recognition of transitive gestures. Pazzaglia *et al.* also observed a close correlation between action execution and recognition in a subgroup of apraxic patients. However, the impairment of recognition in their sample of patients was correlated with lesions at the level of the left inferior frontal gyrus, not of the inferior parietal lobe. The authors of these studies have argued that the discrepancies in their main findings were probably due to different task structures. Pazzaglia *et al.* thus proposed that their recognition task required judgement of the ultimate goal of transitive gestures (or the symbolic meaning of intransitive gestures), while Buxbaum *et al.*'s experiment relied more on identifying kinematic cues. In addition, Buxbaum & Kalenine [77] suggested that the response in Pazzaglia *et al.*'s action-recognition task might have been based on structural rather than functional cues. While this question remains open, these two studies confirm the involvement of the left inferior parietal and ventral premotor cortices in action recognition, possibly at different levels, as well as a tight, although not absolute, functional relationship between action recognition and imitation.

Aphasia has also been known for a long time to induce deficits in the recognition of symbolic gestures and pantomimes in some patients [95]. The question of the comprehension of non-verbal signals by aphasic patients raises an important theoretical issue: is aphasia an impairment specific to the linguistic domain, or is it due to a more general cognitive disorder affecting the use of symbols and signs (asymbolia)? The asymbolia hypothesis has been supported by the experimental work of Duffy & Duffy [96], which showed strong

correlations between scores on pantomime execution and recognition, and language tasks. However, another study concluded that the deficit in pantomiming observed in some aphasic patients might be due to associated apraxia rather than to asymbolia [97]. Aphasic patients are indeed impaired in the comprehension of pantomime in comparison to healthy subjects, but with some dissociation (review in [98]). The left, predominantly frontal, localization of lesions impairing action recognition was confirmed by Tranel *et al.* [99]. A recent controlled study by Saygin *et al.* [100] compared the recognition of actions described by either linguistic (written sentences) or non-linguistic (pictures) cues. Aphasic patients tended to show deficits in both domains, but they were more impaired with linguistic cues, and were also more sensitive to semantic distracters. The authors thus rejected the interpretation of aphasia as fully caused by asymbolia, but nonetheless acknowledged the existence of strong but variable links between linguistic and non-linguistic processes involved in action recognition. Furthermore, the impairment of action recognition in these patients was associated with lesions involving the left inferior frontal areas, in line with Tranel *et al.*'s conclusion. The involvement of the left inferior frontal gyrus in action recognition has also been shown in other tasks. For example, Fazio *et al.* [101] reported that Broca's aphasic patients, though not apraxic, had specific impairment in action and tool naming with respect to object naming, thus supporting the idea that frontal regions might be crucial for action and tool recognition [102]. This specific deficit underlines the double competence of Broca's region, which is not only a language area relating to various aspects and levels of language, but is also a part of the premotor cortex, and as such, is involved in action representation [103]. This consideration has to be regarded in the actual context of embodied language comprehension. Indeed, the processing of action verbs describing leg, mouth or hand movements has been reported to activate motor and premotor areas in a somatotopic manner [104], and may interfere with or facilitate movement execution [105]. These findings suggest that cortical motor regions are involved in action word representation.

(b) *Gesture imitation and speech repetition*
Defective imitation of meaningful or meaningless gestures has often been considered a distinctive sign of apraxia [71,73], and has thus been studied quite extensively. The observation that apraxic patients may be impaired for the imitation of meaningless gestures while being able to reproduce meaningful ones flawlessly [106] has prompted researchers to investigate the processes underlying the imitation of both types of gestures. It has thus been proposed that gesture imitation may be subserved by two distinct routes: a first semantic and indirect route, and a second direct and non-semantic route [106,107]. The former is thought to support imitation of meaningful gestures, while the latter would allow imitation of meaningless gestures by matching the perceived action to the appropriate motor plans. The direct route, however, might subserve the imitation of meaningful gestures in case of damage

to the indirect route. In his seminal work on the topic, Goldenberg and co-workers [106,108] showed that apraxic patients impaired in the imitation of meaningless gestures also showed a deficit in matching the experimenter's posture on a manikin. According to Goldenberg, this demonstrated that the transposition by the direct route of an observed posture into a motor scheme requires the movement to be coded on the basis of a general knowledge of the human body structure. A deficit in imitation of meaningless gestures would thus result from the disturbance of this structural body knowledge, a conceptual body representation that would be independent of the body involved in reproducing a movement (i.e. the subject's, the examiner's or a manikin). This representation is probably supported by the left inferior parietal lobule, which was selectively damaged in Goldenberg and Hagmann's patients [106].

Recently, Schwoebel *et al.* [109] sought to further investigate the involvement of different types of body representation in meaningful and meaningless gesture imitation. Scores on tasks evaluating semantic body knowledge and the body schema (i.e. a dynamic representation of the current relative position of body parts for guiding actions) strongly predicted left brain-damaged patients' performance on imitation and production of meaningful gestures. In contrast, imitation of meaningless gestures depended only on the body schema. These findings confirmed the preferential use of a semantic route for the imitation of meaningful gestures, and the existence of a direct route bypassing semantic knowledge for the imitation of meaningless gestures. Taken together, these observations suggest that imitation of meaningless gestures is more complex than a direct matching between bodies, and is likely to involve both dynamic and more abstract representations of the body.

Gesture imitation is not often evaluated in aphasic patients. However, in the linguistic domain, speech repetition may be conceived of as an equivalent to imitation for manual gestures. According to this idea, speech repetition would be an auditory rather than a visuomotor form of imitation. As mentioned earlier, the idea that language perception relies on audio-motor decoding is not recent and has been defended by Liberman & Mattingly [27] in their motor theory of speech perception. Recent experimental data seem to confirm the existence of a motor resonance of the phonemic percept [110,111]. Speech repetition is often impaired in aphasia, in particular in the case of conduction aphasia. Interestingly, for the purpose of comparing the processes involved in the control of gestures and language, conduction aphasia seems to be associated with lesions of the supramarginal gyrus and the neighbouring planum temporale [65,66], a region also thought to be critically involved in gesture imitation. In addition to impaired repetition, patients with conduction aphasia often exhibit a particular behaviour known as '*conduite d'approche*', characterized by repeated attempts to get closer and closer to the correct utterance. The errors made by these patients are mostly phonemic paraphasias (sound-based speech errors) in which articulators are erroneously selected (e.g. 'basecall' for 'baseball', the /c/ being posterior with respect to the anterior /b/ in terms of

the articulators involved). This is similar to the difficulties seen in patients with ideomotor apraxia when trying to match the position of their hand with respect to other body parts to that demonstrated by the experimenter [106]. This parallel between imitation and repetition fits very well with the case described by Ochipa *et al.* [107], of a patient with a lesion restricted to the inferior parietal lobule and the posterior superior temporal cortex, who exhibited conduction aphasia and apraxia with a particular deficit for imitating tool-use pantomimes. Based on the fact that gesture recognition was preserved in this patient, as speech comprehension usually is in patients with conduction aphasia, the authors even proposed to term this deficit 'conduction apraxia'. Recently, strong support in favour of common functional and anatomical bases for repetition and imitation came from an investigation in patients suffering from primary progressive aphasia who often show various degrees of impairment in different aspects of language and praxis. In their study, Nelissen *et al.* [112] showed that their patients' deficit in speech repetition correlated strongly with their impairment of gesture imitation and discrimination. Lesions in the left rostral inferior parietal lobe, extending to the posterior superior temporal cortex, were significantly associated with these combined impairments. Further, tractography analyses showed that the region most often involved in the lesion was the relay for indirect connections between the superior temporal cortex and the inferior frontal gyrus, offering convincing evidence for a shared neural substrate for gesture imitation and speech repetition and a central role of the left inferior parietal cortex in these abilities.

(c) Pantomime and gestural communication

In apraxic patients, meaningful intransitive gestures have been much less studied than transitive gestures involving object or tool use. This may appear paradoxical given that intransitive gestures, as they are commonly tested, are in fact symbolic gestures (e.g. waving goodbye) strongly related to gestural expression and thus potentially linked to language. However, as argued by Goldenberg *et al.* [113], pantomimes of transitive gestures are also of interest for the present purpose, as they also constitute a link between tool use and communicative manual actions. Indeed, these gestures symbolize the tool and the associated action, and may be used to communicate or to demonstrate proper use of the tool. Pantomimes may thus have been essential in the development of human technology and of a gestural language.

In the context of examining the links between the cerebral control of gestures and language, it is interesting to note that the concepts of transitivity and intransitivity also apply in language to verb argument structure. Indeed, verbs can be differentiated as a function of the number of arguments they require. Intransitive verbs only need an agent, while transitive verbs need an agent and an object. Verbs can even be ditransitive, and require an agent, an object and a recipient. Broca's aphasic patients, whose linguistic production is agrammatic, tend to produce simple rather than complex verb argument structure [114], thus favouring intransitive rather than transitive verbs, as apraxic patients do with gestures (see below). Brain imaging investigations of the neural network underlying the processing of verb argument structure have highlighted not only the role of anterior language areas (i.e. the inferior frontal gyrus), but have also put forward the decisive role played by the parietal cortex, and especially the angular gyrus [115,116].

In line with classical reports of studies in apraxic patients, Mozaz *et al.* [117] showed that apraxic patients are less impaired when performing intransitive than transitive gestures. This was later confirmed by Buxbaum *et al.* [118], who found, in addition, a much weaker relation between imitation and recognition of intransitive gestures. In agreement with this, Heath *et al.* [119] found that a similar percentage of patients with right or left hemispheric damage were impaired for performing meaningful intransitive gestures, suggesting that these gestures are neither unequivocally linked to apraxia nor strongly lateralized. Impairment of tool-use pantomimes would thus be more specific of apraxia than that of intransitive gestures. However, recent studies have challenged the classical view of distinct anatomo-functional bases for the production of transitive and intransitive gestures. Instead, these reports [117–119] suggest that both categories of gestures might rely on the same mechanisms, with transitive gestures being simply more difficult to perform than intransitive ones. Tool-use pantomime is a particularly complicated task since it requires motor imagery and cognitive analysis of the gesture before producing it in detail. In contrast, actual tool use may be guided by the structure of the object itself (affordances) as well as sensory information during hand–object interaction. Accordingly, as pointed out by Carmo & Rumiati [120], no double dissociation has been found between the performance of transitive and intransitive gestures in left-brain-damaged patients: while some patients have been described with impaired transitive and preserved intransitive gestural performance, the reverse profile has, to our knowledge, never been reported (see Stamenova *et al.* [121] for cases in right-brain-damaged patients). Carmo & Rumiati thus analysed the performance of healthy individuals on an imitation task involving transitive and intransitive gestures, and found that they were better at imitating intransitive than transitive movements, in accordance with apraxic patients' difficulties with transitive gestures. In keeping with this idea, Frey [26] and Kroliczak & Frey [122] observed that, in healthy individuals, transitive and intransitive gestures activate the same, hand-independent network in the left hemisphere, suggesting indeed that the same mechanisms might be at play in both conditions. Regarding the neural substrate for the ability to pantomime tool-use actions, recent findings challenged the long-standing notion that pantomimes were primarily supported by the left inferior parietal cortex, thought to store praxic representations [41,72,93]. Indeed, a recent study in apraxic patients showed, using current lesion reconstruction and mapping techniques, that the critical region for the ability to pantomime tool-use actions is rather the posterior part of the left inferior frontal gyrus [113].

In sum, the ability to pantomime, which constitutes a link between manual tool use and communication, is very often disrupted in apraxia, and seems to rely mainly on the integrity of the left inferior frontal gyrus. The impact of brain damage and apraxia on intransitive communicative gestures, however, requires further investigation. While neuroimaging studies in healthy individuals suggest that they are supported by the same neural substrate, recent neuropsychological reports suggest that intransitive, symbolic gestures might be less tied to left hemisphere function than transitive gestures [119,121]. In particular, their relation to genuine gestural communication, language in general, and tool-use gestures, still needs to be explored.

Gestural communication, and the link between language and gestures, has been more largely studied in aphasic patients. These studies have considered several classes of communicative gestures, in contrast to the specific case of pantomime illustrated above and extensively examined in apraxic patients. McNeill [123] proposed a classification of these communicative gestures, organized along a continuum. He distinguishes co-speech gestures, spontaneously used during communication, 'language-like gestures' (grammatically integrated into the utterance), pantomimes (where speech is not necessary), emblems (which have a standard of well-formed-ness, like the sign 'ok') and finally sign languages used by the deaf. Along this continuum, idiosyncratic gestures are progressively replaced by socially regulated signs, the obligatory presence of speech declines (i.e. co-speech gestures accompany spoken language but are not sufficient to convey meaning by themselves, in contrast to sign languages), while language properties embedded in gestures increase. On the contrary to co-speech gestures, sign languages have genuine linguistic properties, with distinctive semantics and syntactic rules, like spoken languages do.

Co-speech gestures are frequent in human communication and have diverse functional roles with large cultural variations [124], but cannot be considered as linguistic gestures by themselves [123]. They are idiosyncratic and individual, and convey meaning by different ways (iconic, metaphoric, deictic, beats, cohesive, etc.) that are radically different from language. First, co-speech gestures are global and synthetic (i.e. neither combinatorial nor hierarchical). Second, they have no standard of form. Third, they lack duality of patterning (in contrast to words where sounds and meanings are both separately structured and arbitrarily linked). However co-speech gestures are intimately linked to language since gestures and speech are synchronous and 'semantically and pragmatically co-expressive'. According to McNeill's hypothesis [123] 'speech and gesture are elements of a single integrated process of utterance formation in which there is a synthesis of opposite modes of thought. Utterances and thought realized in them are both imagery and language'. Regarding the impact of aphasia on co-speech gestures, while it is recognized that aphasic patients may spontaneously use them, there is still no agreement on their level of gestural impairment relative to the level of verbal impairment. For some authors, gestural and verbal expressions are both impaired, owing to a common deficit in communication [125,126]. Other studies claim that gestural expression is less impaired than language, or even that it is more developed than in healthy individuals, perhaps as a result of compensation [46,127]. The neural bases of expressive gestures in healthy individuals have attracted much attention in recent years [128]. However, little is known about the control of expressive gestures in aphasic patients. Several clinical trials have analysed the use of gestures for the rehabilitation of aphasic patients but the results are still unclear [129–133].

At the opposite end of McNeill's continuum, other studies have investigated the impact of brain lesions on the ability to sign. Poizner *et al.* [134,135] observed deaf signers who became aphasic for sign language. Importantly, the impairment was specific to the linguistic components of sign language and dissociated from the production or recognition of non-linguistic gestures and the general ability to use symbols. Studies using functional neuroimaging in neurologically intact deaf signers demonstrated that the neural systems supporting signed language were lateralized to the left hemisphere and very similar to the systems supporting spoken languages, with the additional involvement of the left parietal lobe [136].

In sum, a direct comparison between the impact of apraxia and aphasia on gestural communication is difficult based on the existing literature. Indeed, genuine co-speech gestures are usually examined only in relation to aphasia, without a clear analysis of the impact of potentially associated apraxia. In addition, the impact of apraxia on intransitive gestures, which are mostly emblems as classically assessed in the clinical examination, needs further investigation. Conflicting data in the literature on the impact of aphasia on communicative gestures may also be due to confusion between different categories of motor behaviours along McNeil's continuum, bearing very different relationships with speech and language. In addition, little is known about the spontaneous use of different kinds of communicative gestures in aphasic and apraxic patients. Despite these limitations, however, some links between gestures and language have been demonstrated. In particular, spoken and signed languages are supported by largely overlapping networks [136] (although they both can be dissociated from the production and recognition of non-linguistic gestures [134]). In addition, pantomime of tool use relies mainly on the brain region encompassing Broca's area [113]. Together, the findings reviewed in this section clearly show that if the networks subserving the various aspects of language and praxis are not identical, they largely overlap, with key nodes in the left inferior frontal, inferior parietal and superior temporal cortices.

Further, several studies suggested functional links between language and praxis, raising the possibility of shared processes between both cognitive abilities. In §5, we will try to provide clues as to whether language and praxis may indeed share some common resources.

5. COMMON RESOURCES FOR PRAXIS AND LANGUAGE

The left cerebral hemisphere is considered to play a dominant role for many aspects of praxic and linguistic

behaviours. It is certainly true that some functions related to praxis (e.g. naturalistic multi-step actions [137]) or language (e.g. pragmatic communication [46]) seem to be supported by both hemispheres, or even to be lateralized to the right hemisphere (e.g. matching of finger postures [138] or prosodic processing [44]). However, most praxic and linguistic processes appear consistently lateralized to the left hemisphere. As reviewed earlier, this is true for phonological, semantic and syntactic processing for speech comprehension and production [47]. As for praxis, the following functions depend on the left hemisphere: pantomime [93], actual tool use [68], gesture imitation and recognition [93,107,112] and conceptual knowledge about action and tools [73]. In sum, while cases of atypical cerebral dominance for praxis and language have been described in the neuropsychological literature [138,139], it remains that aphasia and apraxia are both caused by left hemispheric lesions in the vast majority of patients [48,71]. With respect to the evolutionary hypotheses outlined in the introduction, it is interesting that the cerebral lateralization for praxis is more strongly linked to the dominance for language than to manual preference [138,140,141]. This might be due to the necessary interactions between praxic representations and other linguistic-related processes, such as semantics and conceptual knowledge [26,76].

Beyond the observation that these symptoms usually arise after lesions to the same hemisphere, it is striking that apraxia and aphasia are very often associated in right-handed patients with left brain damage [71]. However, the frequent co-occurrence and common hemispheric lateralization are not sufficient to conclude that aphasia and apraxia reflect the same impairment. For example, apraxic patients may exhibit deficits that are linked to non-linguistic processes, such as mechanical reasoning. Indeed, they often have difficulties in solving mechanical puzzles, which require inferring the function of a tool or of an object solely from its structure [142], or in technical reasoning [143]. Thus, praxis implies some left lateralized cognitive ability important for actual tool use but independent from linguistic capacity.

The frequent association of aphasia and apraxia is often seen as the mere consequence of the fact that the cortical regions mediating language and praxis overlap and are vascularized by a common arterial blood supply; thus, there is a high probability that they will both be damaged in the case of stroke. The fact that the co-occurrence of aphasia and apraxia is almost systematic [144] has brought some support to this conception. A clinical study specifically aimed at evaluating the frequency of the co-occurrence of apraxia and aphasia in a large sample of left-brain-damaged right-handed patients indeed reported the existence of a double dissociation between these two disorders in a minority of cases: of 699 patients, 10 had apraxia without aphasia, and 149 were aphasic but not apraxic [42]. In neuropsychology, the existence of a double dissociation between two disorders is usually considered as evidence for a functional independence of the two corresponding cognitive functions (e.g. [145]). However, as argued by Iacoboni & Wilson [146], it is well known that cerebral organization

shows large inter-individual variability at many levels. It is thus possible that the minority of patients showing this double dissociation between aphasia and apraxia, especially in such low proportions as for apraxia without aphasia, may rather represent the two tails of the probabilistic distribution of inter-individual variability for the anatomo-functional organization of language and praxis systems. According to this view, a large majority of individuals would actually have shared neural networks for both abilities. Other interpretations of the frequent association of aphasia and apraxia have thus proposed that both disorders reflect the disturbance of common mechanisms, which may be conceived, for example, as a global communicative or semantic competence [96], or as a left hemisphere specialization for the control of complex sequences [2,124,147]. While it seems unlikely, in light of the literature reviewed here, that apraxia and aphasia strictly reflect a common disorder, many findings coming from neuropsychology and other fields suggest that language and praxis networks may actually intersect and share some common processes.

In particular, the motor aspects of speech and praxis, especially their requirements for sequentially selecting and combining successively different effectors, have long been considered to be potentially underpinned by a common specialization of the left hemisphere for such behaviours [147]. Furthermore, as we have discussed previously, speech repetition and gesture imitation, in addition to bearing similarities as gestural or linguistic imitative behaviours, seem to share common anatomo-functional bases. Indeed, the left inferior parietal cortex, and in particular the supramarginal gyrus, is critical for gesture imitation [79,107,112,148] as well as for repetition [65,66,112]. In addition, functional magnetic resonance imaging studies have allowed a network subserving audio-motor transformations and phonological processing, which are necessary for speech production and repetition, to be delineated [47,149]. This network links the anterior part of Broca's region (see also Kotz *et al.* [31] for the involvement of Broca's area in phonology perception) to the posterior part of the planum temporale and the supramarginal gyrus. A similar network has also been implicated in gesture imitation [113]. In line with these observations, a common underlying mechanism for repetition and imitation has recently been proposed by Iacoboni & Wilson [146]. In their model, the left inferior parietal cortex is thought to have the critical role of reinforcing associations between the appropriate forward and inverse models for language and gestures perception and production. Inverse modelling, allowing the translation of perceived speech or actions into motor plans, is implemented by connections (via inferior parietal areas) between the superior temporal cortex (which encompasses Wernicke's area), involved in the perception of speech and actions, and ventral premotor areas (including Broca's region), which support motor planning and programming for speech and gesture production. Forward modelling, on the other hand, allows the sensory consequences of the intended motor acts, critical for online motor control, to be predicted. These forward models are thought to be implemented by projections, again via the inferior parietal cortex, from the ventral premotor cortex to superior temporal

areas. In this framework, damage to the inferior parietal lobule would thus cause difficulties in updating the inverse model based on the forward model, resulting in impaired repetition and imitation.

Furthermore, language and praxis also seem to interact strongly at the semantic level. Recent studies have shown that the semantic system is much more distributed than originally thought. Binder *et al.* [150] carried out a meta-analysis of functional neuroimaging studies on semantic processing in healthy individuals. They concluded that the semantic system clearly depends on large networks distributed in the temporal, frontal and parietal cortices, predominantly, but not exclusively, in the left hemisphere. This network is constituted of heteromodal association areas, similar to the 'default network[1]', but with little overlap with the distributed network activated by sensorimotor activity. Valuable insight into the cerebral bases of semantic processing has also been provided by neuropsychological investigations in brain-damaged patients. These studies have shown that some patients with very focal brain lesions may present with selective impairment of naming objects of specific semantic categories (e.g. living things versus inanimate objects), thus allowing inferences about the precise semantic function supported by the damaged area. In particular, a double dissociation has been shown between the capacity to name verbs or nouns ([151], review in [152]). Interestingly, while naming nouns involves cortical areas closer to the regions activated by object recognition tasks, naming verbs is supported by areas closer to the frontal motor regions [153]. This suggests that the motor system may be involved in action representation, which could serve action recognition with the aim of pantomiming or imitating, or with the aim of denominating [34]. In this context, a recent study in healthy subjects showed that symbolic gestures and spoken language activated the same left lateralized network, corresponding to Broca's & Wernicke's areas. According to the authors, this suggests that this system 'is not committed to language processing but may function as a modality-independent semiotic system that plays a broader role in human communication, linking meaning to symbols' [154]. In line with this, MacSweeney *et al.* [136] also identified this network, with the addition of the left inferior parietal cortex, as the neural substrate for signed language in deaf individuals.

Finally, as suggested earlier, another potential functional relation between praxis and language lies in the hierarchical organization of those behaviours. Is syntax exclusive to the linguistic domain, or are complex actions and music endowed with hierarchical rules akin to linguistic syntax [155,156]? In the domain of language, Broca's aphasia is classically qualified as 'agrammatic' owing to impairment in producing grammatical sentences and in processing syntactic markers. Accordingly, a wealth of brain imaging studies has reported the activation of the caudal part of Broca's area (Broadmann area 44) in tasks involving syntactic processing [85–87]. In accordance with these observations, Grodzinsky & Santi [157] proposed that abstract linguistic abilities are neurologically coded, and that Broca's area might play a specific role in syntax processing. However, other authors have

proposed that Broca's aphasics' impairment in syntactic processing might rather be due to a lessening of available resources [56], and that the role of Broca's area might instead be to bind together the semantic, syntactic and phonological levels of language [158]. In the praxic domain, on the other hand, complex actions can be conceived of as structured according to three levels of organization: hierarchical (goals and subgoals), temporal (action sequences) and spatial (embodiment of tools). This structure could be paralleled with the organization of language [34,38]. Interestingly, the hierarchical control of action appears to involve Broca's region and its right homologue [159,160]. This is consistent with the fact that planning deficits, characteristic of dysexecutive syndromes, are attributed to lesions of the frontal lobes [161] and that both complex sequential actions (e.g. preparing a cup of coffee) [137] and pragmatic communication in a social context [46] seem to depend on both hemispheres. In addition, there is evidence for a convergence between language and praxis at the syntactic level as well. For example, Fazio *et al.* [101] have recently provided support for the existence of a link, potentially supported by Broca's area, between language syntax and sequential organization of observed actions. These authors have demonstrated that agrammatic patients with lesions involving Broca's area are also impaired in a non-linguistic test consisting of ordering action sequences. This functional relation between action recognition and language perception has recently been confirmed and further characterized by Sitnikova *et al.* [162], who examined event-related potentials in an action structural violation paradigm. In this study, healthy participants were presented with movies depicting everyday familiar tasks involving the use of tools. The authors found that the introduction of a tool irrelevant to the action context (e.g. an iron in the context of cutting bread) elicited a neurophysiological response usually linked to syntactic processes and violation detection. Interestingly, the stimuli used in this paradigm are highly reminiscent of the errors made by some patients with conceptual apraxia. These patients may indeed be unable to judge the appropriateness of the gesture demonstrated by an experimenter in association with a given tool or object [78], or they may also make similar errors when asked to demonstrate the gestures themselves, either choosing the wrong tool in a given context or executing the wrong action in response to a visually presented object [73].

6. CONCLUSION

Several hypotheses on the emergence of human culture and cognitive capacities have proposed a close evolutionary link between praxis, including tool use, and language, which are uniquely developed in the human species. Some of them have in addition proposed that gestural communication, involving action recognition and imitation, might have constituted an intermediary stage between the development of tool manufacture and use, and the emergence of spoken language. These hypotheses predict a strong relationship between the neural substrates and cognitive processes involved in language and praxis. Here we have addressed this

question by reviewing the neuropsychological literature pertaining to the impairments of language and praxis after brain lesions, respectively termed aphasia and apraxia. In particular, we examined their impact on action recognition, imitation and communicative gestures, as well as the possible anatomo-functional links between the two neural systems supporting the two uniquely human cognitive abilities.

Research in brain-damaged patients as well as healthy individuals has shown that the functional anatomy of language and praxis is complex and organized in several networks, mainly lateralized to the left hemisphere. These praxic and linguistic networks partly overlap, with critical nodes located in the superior temporal, rostral inferior parietal and ventral premotor cortices. In contrast to what has often been argued, the frequent co-occurrence of aphasia and apraxia in left-brain-damaged patients may not be only the mere consequence of the proximity of the cortical areas involved. Rather, this phenomenon might reflect the fact that praxis and language networks actually intersect and share some common functional processes. This view is compatible with the existence of joint linguistic and praxic impairments, reflecting common deficient mechanisms, as well as with the dissociation between other manifestations of aphasia and apraxia. One demonstrative example of dissociation is the case of aphasic patients who previously used sign language and have selective impairment of the linguistic aspects of their gestures [134–136].

Accordingly, several recent studies in left-brain-damaged patients have suggested strong links between speech repetition and gesture imitation, which may involve common neural substrates and mechanisms, with a critical role played by the left inferior parietal cortex. Both aphasia and apraxia may induce impairments of action recognition, but some evidence suggests that the mechanisms involved are at least partially dissociable. Regardless, investigations in aphasic and apraxic patients suggest involvement of the left inferior frontal gyrus, including Broca's area, as well as the left inferior parietal lobe. As for gestural communication, the available literature does not allow comparison of the consequences of the aphasia and apraxia, since different categories of gestures have been studied in relation to each syndrome. This question will thus need further and more specific investigation. The available literature leads to the conclusion that the hypothesis of asymbolia is too general to explain the complex relations between gesture communication, aphasia and apraxia [100,163,164]. However, evidence from neuropsychological and neuroimaging studies converges to suggest that tool-use pantomimes, which may have been critical for the transmission and propagation of human tool manufacture and use, depend at least partly on the same neural network as actual tool use and symbolic, intransitive gestures. In particular, lesions involving the left inferior frontal gyrus, known to be strongly involved in speech production and comprehension, seem to be especially associated with impairment of tool-use pantomimes. In sum, neuropsychological studies of linguistic and praxic disorders show that both systems interact more or less depending on the context and on which aspects these complex cognitive behaviours are considered. Mounting evidence suggests that the phonological,

semantic and syntactic levels of language share some common cognitive resources with the praxis system. This is consistent with the hypothesis of common phylogenetic and ontogenetic origins for language and praxis [9]. However, neuropsychological data do not allow confirmation or rejection of the hypothesis of an intermediary stage of gestural communication between the development of tool use and the emergence of language.

In addition to the development of long-range interconnected networks, evolution within some strategic areas in the left hemisphere might have conditioned the appearance and lateralization of complex human behaviour. The left lateralization might be attributed to an asymmetry of the columnar micro-architecture of the cortex inducing an asymmetry of some general processes then leading to a differential development of functions [165]. Along this line of reasoning, Goldenberg [67] proposes that the specific role of the left parietal lobe is based on categorical apprehension of spatial relationships, consistently with the left hemisphere preference for categorical coding (by opposition to coordinate coding). Similarly, the hemispheric lateralization for speech could result from an asymmetry of cortical temporal tuning, itself inducing an asymmetry of audio-motor processes [166]. According to this hypothesis, the left hemisphere might be specialized for the perception and production of sounds in the 28–40 Hz frequency domain (i.e. perception of phonemes and tongue movements) while the right hemisphere might be specialized in the 3–6 Hz frequency domain (i.e. syllabic rate and jaw movements). The role of Broca's area (or more generally speaking, the left inferior frontal gyrus) is now being revisited and ardently disputed [157]. In addition to its contribution to the human mirror system [167], it could have a generic function for hierarchical processing and nesting of chunks and sequences [160], unification of the different aspects of language [158] or binding meaning and symbol [154]. This kind of generic function might be a common resource for action and language, grounded in the left hemisphere and acting like a node in a complex and bilateral distributed network. These processes probably condition the richness and complexity of human activity.

ENDNOTE
[1] The default network is activated in resting states for specific tasks 'interrupting the stream of consciousness'.

REFERENCES
1 Annett, M. 2006 The distribution of handedness in chimpanzees: estimating right shift in Hopkins' sample. *Laterality* **11**, 101–109.

2 Kimura, D. & Archibald, Y. 1974 Motor functions of the left hemisphere. *Brain* **97**, 337–350. (doi:10.1093/brain/97.1.337)

3 Corbetta, D. 2003 Right-handedness may have come first: evidence from studies in human infants and non-human primates. *Behav. Brain Sci.* **26**, 217–218. (doi:10.1017/S0140525X03320060)

4 Bradshaw, J. L. & Nettleton, N. C. 1982 Language lateralization to the dominant hemisphere: tool use, gesture and language in hominid evolution. *Curr. Psychol. Rev.* **2**, 171–192. (doi:10.1007/BF02684498)

5 Corballis, M. C. 2003 From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* **26**, 199–208. (doi:10.1017/S0140525X03000062)

6 Ambrose, S. H. Paleolithic technology and human evolution. *Science* **291**, 1748–1753. (doi:10.1126/science.1059487)

7 Tomasello, M. 1999 The human adaptation for culture. *Annu. Rev. Anthropol.* **28**, 509–529. (doi:10.1146/annurev.anthro.28.1.509)

8 Steele, J. & Uomini, N. 2009 Can the archaeology of manual specialization tell us anything about language evolution? A survey of the state of play. *Steps to a (neuro-) archaeology of mind* (eds L. Malafouris & C. Renfrew). *Camb. Archaeol. J.* **19**, 97–110.

9 Greenfield, P. M. 1991 Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* **14**, 531–551. (doi:10.1017/S0140525X00071235)

10 Iriki, A., Tanaka, M. & Iwamura, Y. 1996 Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* **7**, 2325–2330. (doi:10.1097/00001756-199610020-00010)

11 Obayashi, S., Suhara, T., Kawabe, K., Okauchi, T., Maeda, J., Akine, Y., Onoe, H. & Iriki, A. 2001 Functional brain mapping of monkey tool use. *Neuroimage* **14**, 853–861. (doi:10.1006/nimg.2001.0878)

12 Frey, S. H. 2007 What puts the how in where? Tool use and the divided visual streams hypothesis. *Cortex* **43**, 368–375. (doi:10.1016/S0010-9452(08)70462-3)

13 Arbib, M. A., Bonaiuto, J. B., Jacobs, S. & Frey, S. H. 2009 Tool use and the distalization of the end-effector. *Psychol. Res.* **73**, 441–462. (doi:10.1007/s00426-009-0242-2)

14 Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A. C. & Farne, A. 2009 Tool-use induces morphological updating of the body schema. *Curr. Biol.* **19**, R478–R479. (doi:10.1016/j.cub.2009.05.009)

15 Ungerleider, L. G. & Mishkin, M. 1982 Two cortical visual system. In *Analysis of visual behaviour* (eds D. L. Ingle, M. A. Goodale & R. J. W. Mansfield), pp. 549–586. Cambridge, MA: MIT Press.

16 Goodale, M. A. & Milner, A. D. 1992 Separate visual pathways for perception and action. *Trends Neurosci.* **15**, 20–25. (doi:10.1016/0166-2236(92)90344-8)

17 Rizzolatti, G. & Luppino, G. 2001 The cortical motor system. *Neuron* **31**, 889–901. (doi:10.1016/S0896-6273(01)00423-8)

18 Jacobs, S., Danielmeier, C. & Frey, S. H. 2010 Human anterior intraparietal and ventral premotor cortices support representations of grasping with the hand or a novel tool. *J. Cogn. Neurosci.* **22**, 2594–2608. (doi:10.1162/jocn.2009.21372)

19 Umilta, M. A., Escola, L., Intskirveli, I., Grammont, F., Rochat, M., Caruana, F., Jezzini, A., Gallese, V. & Rizzolatti, G. 2008 When pliers become fingers in the monkey motor system. *Proc. Natl Acad. Sci. USA.* **105**, 2209–2213. (doi:10.1073/pnas.0705985105)

20 Johnson-Frey, S. H. 2003 What's so special about human tool use? *Neuron* **39**, 201–204. (doi:10.1016/S0896-6273(03)00424-0)

21 Hodges, J. R., Bozeat, S., Lambon Ralph, M. A., Patterson, K. & Spatt, J. 2000 The role of conceptual knowledge in object use evidence from semantic dementia. *Brain* **123**, 1913–1925. (doi:10.1093/brain/123.9.1913)

22 Johnson-Frey, S. H. 2004 The neural bases of complex tool use in humans. *Trends Cogn. Sci.* **8**, 71–78. (doi:10.1016/j.tics.2003.12.002)

23 Lewis, J. W. 2006 Cortical networks related to human use of tools. *Neuroscientist* **12**, 211–231. (doi:10.1177/1073858406288327)

24 Creem, S. H. & Proffitt, D. R. 2001 Grasping objects by their handles: a necessary interaction between cognition and action. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 218–228. (doi:10.1037/0096-1523.27.1.218)

25 Randerath, J., Li, Y., Goldenberg, G. & Hermsdorfer, J. 2009 Grasping tools: effects of task and apraxia. *Neuropsychologia* **47**, 497–505. (doi:10.1016/j.neuropsychologia.2008.10.005)

26 Frey, S. H. 2008 Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Phil. Trans. R. Soc. B* **363**, 1951–1957. (doi:10.1098/rstb.2008.0008)

27 Liberman, A. M. & Mattingly, I. G. 1985 The motor theory of speech perception revised. *Cognition* **21**, 1–36. (doi:10.1016/0010-0277(85)90021-6)

28 Rizzolatti, G. & Craighero, L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)

29 Fogassi, L. & Luppino, G. 2005 Motor functions of the parietal lobe. *Curr. Opin. Neurobiol.* **15**, 626–631. (doi:10.1016/j.conb.2005.10.015)

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

31 Kotz, S. A., D'Ausilio, A., Raettig, T., Begliomini, C., Craighero, L., Fabbri-Destro, M., Zingales, C., Haggard, P. & Fadiga, L. 2010 Lexicality drives audio-motor transformations in Broca's area. *Brain Lang.* **112**, 3–11. (doi:10.1016/j.bandl.2009.07.008)

32 Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. 1995 Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* **73**, 2608–2611.

33 Rizzolatti, G., Fogassi, L. & Gallese, V. 2001 Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670. (doi:10.1038/35090060)

34 Pulvermuller, F. & Fadiga, L. 2010 Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* **11**, 351–360. (doi:10.1038/nrn2811)

35 Hickok, G. 2009 Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* **21**, 1229–1243. (doi:10.1162/jocn.2009.21189)

36 Toni, I., de Lange, F. P., Noordzij, M. L. & Hagoort, P. 2008 Language beyond action. *J. Physiol. Paris* **102**, 71–79. (doi:10.1016/j.jphysparis.2008.03.005)

37 Rizzolatti, G. & Arbib, M. A. 1998 Language within our grasp. *Trends Neurosci.* **21**, 188–194. (doi:10.1016/S0166-2236(98)01260-0)

38 Arbib, M. A. 2010 Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain Lang.* **112**, 12–24. (doi:10.1016/j.bandl.2009.10.001)

39 Arbib, M. A. 2006 Aphasia, apraxia and the evolution of the language-ready brain. *Aphasiology* **20**, 1125–1155. (doi:10.1080/02687030600741683)

40 Broca, P. 1861 Remarques sur le siège de la faculté du langage articulé suivi d'une observation d'aphémie (perte de la parole). *Bull. Soc. Anat.* **6**, 330–357.

41 Liepmann, H. 1908 *Drei aufsatze aus dem apraxiegebiet.* Berlin, Germany: Karger.

42 Papagno, C., Della Sala, S. & Basso, A. 1993 Ideomotor apraxia without aphasia and aphasia without apraxia: the anatomical support for a double dissociation. *J. Neurol. Neurosurg. Psychiat.* **56**, 286–289. (doi:10.1136/jnnp.56.3.286)

43 Benson, F. & Ardila, A. 1996 *Aphasia: a clinical perspective.* New York, NY: Oxford University Press.

44 Ross, E. D. & Monnot, M. 2008 Neurology of affective prosody and its functional-anatomic organization in right hemisphere. *Brain Lang.* **104**, 51–74. (doi:10.1016/j.bandl.2007.04.007)

45 Sammler, D., Kotz, S. A., Eckstein, K., Ott, D. V. & Friederici, A. D. 2010 Prosody meets syntax: the role of the corpus callosum. *Brain* **133**, 2643–2655. (doi:10.1093/brain/awq231)

46 Rousseaux, M., Daveluy, W. & Kozlowski, O. 2010 Communication in conversation in stroke patients. *J. Neurol.* **257**, 1099–1107. (doi:10.1007/s00415-010-5469-8)

47 Vigneau, M., Beaucousin, V., Herve, P. Y., Duffau, H., Crivello, F., Houde, O., Mazoyer, B. & Tzourio-Mazoyer, N. 2006 Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* **30**, 1414–1432. (doi:10.1016/j.neuroimage.2005.11.002)

48 Caplan, D. 2003 Aphasic syndromes. In *Clinical neuropsychology* (eds K. M. Heilman & E. E. Valenstein), pp. 14–34, 4th edn. Oxford, UK: Oxford University Press.

49 Noppeney, U. & Wallesch, C. W. 2000 Language and cognition: Kurt Goldstein's theory of semantics. *Brain Cogn.* **44**, 367–386. (doi:10.1006/brcg.1999.1199)

50 Geschwind, N. 1965 Disconnexion syndromes in animals and man. I. *Brain* **88**, 237–294. (doi:10.1093/brain/88.2.237)

51 Geschwind, N. 1965 Disconnexion syndromes in animals and man. II. *Brain* **88**, 585–644. (doi:10.1093/brain/88.3.585)

52 Catani, M. & ffytche, D. H. 2005 The rises and falls of disconnection syndromes. *Brain* **128**, 2224–2239. (doi:10.1093/brain/awh622)

53 Hillis, A. E. 2007 Aphasia: progress in the last quarter of a century. *Neurology* **69**, 200–213. (doi:10.1212/01.wnl.0000265600.69385.6f)

54 Caramazza, A. & Zurif, E. B. 1976 Dissociation of algorithmic and heuristic processes in language comprehension: evidence from aphasia. *Brain Lang.* **3**, 572–582. (doi:10.1016/0093-934X(76)90048-1)

55 O'Grady, W. & Lee, M. 2001 The isomorphic mapping hypothesis: evidence from Korea. *Brain Cogn.* **46**, 226–230. (doi:10.1016/S0278-2626(01)80072-5)

56 Caplan, D. 2006 Aphasic deficits in syntactic processing. *Cortex* **42**, 797–804. (doi:10.1016/S0010-9452(08)70420-9)

57 Caplan, D., Baker, C. & Dehaut, F. 1985 Syntactic determinants of sentence comprehension in aphasia. *Cognition* **21**, 117–175. (doi:10.1016/0010-0277(85)90048-4)

58 Ogar, J., Slama, H., Dronkers, N., Amici, S. & Gorno-Tempini, M. L. 2005 Apraxia of speech: an overview. *Neurocase* **11**, 427–432. (doi:10.1080/13554790500263529)

59 Ogar, J., Willock, S., Baldo, J., Wilkins, D., Ludy, C. & Dronkers, N. 2006 Clinical and anatomical correlates of apraxia of speech. *Brain Lang.* **97**, 343–350. (doi:10.1016/j.bandl.2006.01.008)

60 Dronkers, N. F. 1996 A new brain region for coordinating speech articulation. *Nature* **384**, 159–161. (doi:10.1038/384159a0)

61 Hillis, A. E., Work, M., Barker, P. B., Jacobs, M. A., Breese, E. L. & Maurer, K. 2004 Re-examining the brain regions crucial for orchestrating speech articulation. *Brain* **127**, 1479–1487. (doi:10.1093/brain/awh172)

62 Ardila, A. 2010 A review of conduction aphasia. *Curr. Neurol. Neurosci. Rep.* **10**, 499–503. (doi:10.1007/s11910-010-0142-2)

63 Luria, A. R. 1976 *Basic problems of neurolinguistics* (trans. B. Haigh). The Hague, The Netherlands: Mouton.

64 Benson, D. F., Sheremata, W. A., Bouchard, R., Segarra, J. M., Price, D. & Geschwind, N. 1973 Conduction aphasia. A clinicopathological study. *Arch. Neurol.* **28**, 339–346.

65 Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M. & Hickok, G. In press. Conduction aphasia, sensory-motor integration, and phonological short-term memory: an aggregate analysis of lesion and fMRI data. *Brain Lang.* (doi:10.1016/j.bandl.2010.12.001)

66 Fridriksson, J., Kjartansson, O., Morgan, P. S., Hjaltason, H., Magnusdottir, S., Bonilha, L. & Rorden, C. 2010 Impaired speech repetition and left parietal lobe damage. *J. Neurosci.* **30**, 11 057–11 061. (doi:10.1523/JNEUROSCI.1120-10.2010)

67 Goldenberg, G. 2009 Apraxia and the parietal lobes. *Neuropsychologia* **47**, 1449–1459. (doi:10.1016/j.neuropsychologia.2008.07.014)

68 Goldenberg, G. & Spatt, J. 2009 The neural basis of tool use. *Brain* **132**, 1645–1655. (doi:10.1093/brain/awp080)

69 Roy, E. A. & Square, P. 1985 Common considerations in the studies on limb, verbal and oral apraxia. In *Neuropsychological studies of apraxia and related disorders* (ed. E. Roy), pp. 111–162. Amsterdam, The Netherlands: North Holland.

70 Heilman, K. M. & Rothi, L. J. 1997 Limb apraxia: a look back. In *Apraxia: the neuropsychology of action* (eds L. J. Rothi & K. M. Heilman), pp. 7–18. Hove, UK: Psychology Press.

71 Heilman, K. M. & Rothi, L. J. 2003 Apraxia. In *Clinical neuropsychology* (eds K. M. Heilman & E. E. Valenstein), pp. 215–236, 4th edn. New York, NY: Oxford University Press.

72 Heilman, K. M., Rothi, L. J. & Valenstein, E. 1982 Two forms of ideomotor apraxia. *Neurology* **32**, 342–346.

73 De Renzi, E. & Lucchelli, F. 1988 Ideational apraxia. *Brain* **3**, 1173–1185. (doi:10.1093/brain/111.5.1173)

74 Trojano, L., Labruna, L. & Grossi, D. 2007 An experimental investigation of the automatic/voluntary dissociation in limb apraxia. *Brain Cogn.* **65**, 169–176. (doi:10.1016/j.bandc.2007.07.010)

75 Leiguarda, R. & Marsden, C. 2000 Limb apraxias: higher-order disorders of sensorimotor integration. *Brain* **123**, 860–879. (doi:10.1093/brain/123.5.860)

76 Buxbaum, L. J. 2001 Ideomotor apraxia: a call to action. *Neurocase* **7**, 445–458. (doi:10.1093/neucas/7.6.445)

77 Buxbaum, L. J. & Kalenine, S. 2010 Action knowledge, visuomotor activation, and embodiment in the two action systems. *Ann. NY Acad. Sci.* **1191**, 201–218. (doi:10.1111/j.1749-6632.2010.05447.x)

78 Rothi, L. J. & Heilman, K. M. 1997 Introduction to limb apraxia. In *Apraxia: the neuropsychology of action* (eds L. J. Rothi & K. M. Heilman), pp. 1–6. Hove, UK: Psychology Press.

79 Haaland, K. Y., Harrington, D. L. & Knight, R. T. 2000 Neural representations of skilled movement. *Brain* **123**, 2306–2313. (doi:10.1093/brain/123.11.2306)

80 Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B. & Agid, Y. 1995 A selective impairment of hand posture for object utilization in apraxia. *Cortex* **31**, 41–55.

81 Buxbaum, L. J., Sirigu, A., Schwartz, M. F. & Klatzky, R. 2003 Cognitive representations of hand posture in ideomotor apraxia. *Neuropsychologia* **41**, 1091–1113. (doi:10.1016/S0028-3932(02)00314-7)

82 Watkins, K. & Paus, T. 2004 Modulation of motor excitability during speech perception: the role of Broca's area. *J. Cogn. Neurosci.* **16**, 978–987. (doi:10.1162/0898929041502616)

83 Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D. & Iacoboni, M. 2007 The essential role of premotor cortex in speech perception. *Curr. Biol.* **17**, 1692–1696. (doi:10.1016/j.cub.2007.08.064)

84 Fadiga, L. & Craighero, L. 2006 Hand actions and speech representation in Broca's area. *Cortex* **42**, 486–490. (doi:10.1016/S0010-9452(08)70383-6)

85 Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F. & Fazio, F. 2001 Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage* **13**, 110–118. (doi:10.1006/nimg.2000.0668)

86 Friederici, A. D., Meyer, M. & von Cramon, D. Y. 2000 Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* **75**, 289–300. (doi:10.1006/brln.2000.2313)

87 Newman, S. D., Just, M. A., Keller, T. A., Roth, J. & Carpenter, P. A. 2003 Differential effects of syntactic and semantic processing on the subregions of Broca's area. *Brain Res. Cogn. Brain Res.* **16**, 297–307. (doi:10.1016/S0926-6410(02)00285-9)

88 Hickok, G. & Poeppel, D. 2000 Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci.* **4**, 131–138. (doi:10.1016/S1364-6613(00)01463-7)

89 Cubelli, R., Marchetti, C., Boscolo, G. & Della Sala, S. 2000 Cognition in action: testing a model of limb apraxia. *Brain Cogn.* **44**, 144–165. (doi:10.1006/brcg.2000.1226)

90 Halsband, U., Schmitt, J., Weyers, M., Binkofski, F., Grutzner, G. & Freund, H. J. 2001 Recognition and imitation of pantomimed motor acts after unilateral parietal and premotor lesions: a perspective on apraxia. *Neuropsychologia* **39**, 200–216. (doi:10.1016/S0028-3932(00)00088-9)

91 Buccino, G. *et al.* 2001 Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* **13**, 400–404.

92 Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F. 1997 Brain activity during observation of actions. Influence of action content and subject's strategy. *Brain* **120**, 1763–1777. (doi:10.1093/brain/120.10.1763)

93 Buxbaum, L. J., Kyle, K. M. & Menon, R. 2005 On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Brain Res. Cogn. Brain Res.* **25**, 226–239. (doi:10.1016/j.cogbrainres.2005.05.014)

94 Pazzaglia, M., Smania, N., Corato, E. & Aglioti, S. M. 2008 Neural underpinnings of gesture discrimination in patients with limb apraxia. *J. Neurosci.* **28**, 3030–3041. (doi:10.1523/JNEUROSCI.5748-07.2008)

95 Gainotti, G. & Lemmo, M. S. 1976 Comprehension of symbolic gestures in aphasia. *Brain Lang.* **3**, 451–460. (doi:10.1016/0093-934X(76)90039-0)

96 Duffy, R. J. & Duffy, J. R. 1981 Three studies of deficits in pantomimic expression and pantomimic recognition in aphasia. *J. Speech Hear. Res.* **24**, 70–84.

97 Wang, L. & Goodglass, H. 1992 Pantomime, praxis, and aphasia. *Brain Lang.* **42**, 402–418. (doi:10.1016/0093-934X(92)90076-Q)

98 Feyereisen, P. & Seron, X. 1982 Nonverbal communication and aphasia: a review. I. Comprehension. *Brain Lang.* **16**, 191–212. (doi:10.1016/0093-934X(82)90083-9)

99 Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H. & Damasio, A. R. 2003 Neural correlates of conceptual knowledge for actions. *Cogn. Neuropsychol.* **20**, 409–432. (doi:10.1080/02643290244000248)

100 Saygin, A. P., Wilson, S. M., Dronkers, N. F. & Bates, E. 2004 Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia* **42**, 1788–1804. (doi:10.1016/j.neuropsychologia.2004.04.016)

101 Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., Calzolari, F., Granieri, E. & Fadiga, L. 2009 Encoding of human action in Broca's area. *Brain* **132**, 1980–1988. (doi:10.1093/brain/awp118)

102 Martin, A., Wiggs, C. L., Ungerleider, L. G. & Haxby, J. V. 1996 Neural correlates of category-specific knowledge. *Nature* **379**, 649–652. (doi:10.1038/379649a0)

103 Fadiga, L., Craighero, L. & Roy, A. C. 2006 Broca's region: a speech area? In *Broca's region* (eds Y. Grodzinsky & K. Amunts). New York, NY: Oxford University Press.

104 Hauk, O., Johnsrude, I. & Pulvermuller, F. 2004 Somatotopic representation of action words in human motor and premotor cortex. *Neuron* **41**, 301–307. (doi:10.1016/S0896-6273(03)00838-9)

105 Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M. & Nazir, T. A. 2006 Cross-talk between language processes and overt motor behavior in the first 200 msec of processing. *J. Cogn. Neurosci.* **18**, 1607–1615. (doi:10.1162/jocn.2006.18.10.1607)

106 Goldenberg, G. & Hagmann, S. 1997 The meaning of meaningless gestures: a study of visuo-imitative apraxia. *Neuropsychologia* **35**, 333–341. (doi:10.1016/S0028-3932(96)00085-1)

107 Ochipa, C., Rothi, L. J. & Heilman, K. M. 1994 Conduction apraxia. *J. Neurol. Neurosurg. Psychiat.* **57**, 1241–1244. (doi:10.1136/jnnp.57.10.1241)

108 Goldenberg, G. 1995 Imitating gestures and manipulating a manikin—the representation of the human body in ideomotor apraxia. *Neuropsychologia* **33**, 63–72. (doi:10.1016/0028-3932(94)00104-W)

109 Schwoebel, J., Buxbaum, L. J. & Coslett, B. H. 2004 Representations of the human body in the production and imitation of complex movements. *Cogn. Neuropsychol.* **21**, 285–298. (doi:10.1080/02643290342000348)

110 Roy, A. C., Craighero, L., Fabbri-Destro, M. & Fadiga, L. 2008 Phonological and lexical motor facilitation during speech listening: a transcranial magnetic stimulation study. *J. Physiol. Paris* **102**, 101–105. (doi:10.1016/j.jphysparis.2008.03.006)

111 D'Ausilio, A., Pulvermuller, F., Salmas, P., Bufalari, I., Begliomini, C. & Fadiga, L. 2009 The motor somatotopy of speech perception. *Curr. Biol.* **19**, 381–385. (doi:10.1016/j.cub.2009.01.017)

112 Nelissen, N., Pazzaglia, M., Vandenbulcke, M., Sunaert, S., Fannes, K., Dupont, P., Aglioti, S. M. & Vandenberghe, R. 2010 Gesture discrimination in primary progressive aphasia: the intersection between gesture and language processing pathways. *J. Neurosci.* **30**, 6334–6341. (doi:10.1523/JNEUROSCI.0321-10.2010)

113 Goldenberg, G., Hermsdorfer, J., Glindemann, R., Rorden, C. & Karnath, H. O. 2007 Pantomime of tool use depends on integrity of left inferior frontal cortex. *Cereb. Cortex* **17**, 2769–2776. (doi:10.1093/cercor/bhm004)

114 Thompson, C. K. 2003 Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. *J. Neurolinguist.* **16**, 151–167. (doi:10.1016/S0911-6044(02)00014-3)

115 den Ouden, D. B., Fix, S., Parrish, T. B. & Thompson, C. K. 2009 Argument structure effects in action verb naming in static and dynamic conditions. *J. Neurolinguist.* **22**, 196–215. (doi:10.1016/j.jneuroling.2008.10.004)

116 Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parrish, T. B., Gitelman, D. R. & Mesulam, M. M. 2007 Neural correlates of verb argument structure processing. *J. Cogn. Neurosci.* **19**, 1753–1767. (doi:10.1162/jocn.2007.19.11.1753)

117 Mozaz, M., Rothi, L. J., Anderson, J. M., Crucian, G. P. & Heilman, K. M. 2002 Postural knowledge of transitive pantomimes and intransitive gestures. *J. Int. Neuropsychol. Soc.* **8**, 958–962. (doi:10.1017/S1355617702870114)

118 Buxbaum, L. J., Kyle, K., Grossman, M. & Coslett, H. B. 2007 Left inferior parietal representations for skilled hand–object interactions: evidence from stroke and corticobasal degeneration. *Cortex* **43**, 411–423. (doi:10.1016/S0010-9452(08)70466-0)

119 Heath, M., Roy, E. A., Westwood, D. & Black, S. E. 2001 Patterns of apraxia associated with the production of intransitive limb gestures following left and right hemisphere stroke. *Brain Cogn.* **46**, 165–169. (doi:10.1016/S0278-2626(01)80057-9)

120 Carmo, J. C. & Rumiati, R. I. 2009 Imitation of transitive and intransitive actions in healthy individuals. *Brain Cogn.* **69**, 460–464. (doi:10.1016/j.bandc.2008.09.007)

121 Stamenova, V., Roy, E. A. & Black, S. E. 2010 Associations and dissociations of transitive and intransitive gestures in left and right hemisphere stroke patients. *Brain Cogn.* **72**, 483–490. (doi:10.1016/j.bandc.2010.01.004)

122 Kroliczak, G. & Frey, S. H. 2009 A common network in the left cerebral hemisphere represents planning of tool use pantomimes and familiar intransitive gestures at the hand-independent level. *Cereb. Cortex* **19**, 2396–2410. (doi:10.1093/cercor/bhn261)

123 McNeill, D. 1992 *Hand and mind: what gestures reveal about thought.* Chicago, IL: University of Chicago Press.

124 Feyereisen, P. & Seron, X. 1982 Nonverbal communication and aphasia: a review. II. Expression. *Brain Lang.* **16**, 213–236. (doi:10.1016/0093-934X(82)90084-0)

125 Cicone, M., Wapner, W., Foldi, N., Zurif, E. & Gardner, H. 1979 The relation between gesture and language in aphasic communication. *Brain Lang.* **8**, 324–349. (doi:10.1016/0093-934X(79)90060-9)

126 McNeill, D. 1985 So you think gestures are nonverbal? *Psychol. Rev.* **92**, 350–371. (doi:10.1037/0033-295X.92.3.350)

127 Herrmann, M., Reichle, T., Lucius-Hoene, G., Wallesch, C. W. & Johannsen-Horbach, H. 1988 Nonverbal communication as a compensatory strategy for severely nonfluent aphasics? A quantitative approach. *Brain Lang.* **33**, 41–54. (doi:10.1016/0093-934X(88)90053-3)

128 Willems, R. M. & Hagoort, P. 2007 Neural evidence for the interplay between language, gesture, and action: a review. *Brain Lang.* **101**, 278–289. (doi:10.1016/j.bandl.2007.03.004)

129 Christopoulou, C. & Bonvillian, J. D. 1985 Sign language, pantomime, and gestural processing in aphasic persons: a review. *J. Commun. Disord.* **18**, 1–20. (doi:10.1016/0021-9924(85)90010-3)

130 Daumuller, M. & Goldenberg, G. 2010 Therapy to improve gestural expression in aphasia: a controlled clinical trial. *Clin. Rehabil.* **24**, 55–65. (doi:10.1177/0269215509343327)

131 Marangolo, P., Bonifazi, S., Tomaiuolo, F., Craighero, L., Coccia, M., Altoe, G., Provinciali, L. & Cantagallo, A. 2010 Improving language without words: first evidence from aphasia. *Neuropsychologia* **48**, 3824–3833. (doi:10.1016/j.neuropsychologia.2010.09.025)

132 Raymer, A. M., Singletary, F., Rodriguez, A., Ciampitti, M., Heilman, K. M. & Rothi, L. J. 2006 Effects of gesture + verbal treatment for noun and verb retrieval in aphasia. *J. Int. Neuropsychol. Soc.* **12**, 867–882. (doi:10.1017/S1355617706061042)

133 Scharp, V. L., Tompkins, C. A. & Iverson, J. M. 2007 Gesture and aphasia: helping hands? *Aphasiology* **21**, 717–725. (doi:10.1080/02687030701192273)

134 Poizner, H., Bellugi, U. & Iragui, V. 1984 Apraxia and aphasia for a visual-gestural language. *Am. J. Physiol.* **246**, R868–R883.

135 Poizner, H., Bellugi, U. & Klima, E. S. 1990 Biological foundations of language: clues from sign language. *Annu. Rev. Neurosci.* **13**, 283–307. (doi:10.1146/annurev.ne.13.030190.001435)

136 MacSweeney, M., Capek, C. M., Campbell, R. & Woll, B. 2008 The signing brain: the neurobiology of sign language. *Trends Cogn. Sci.* **12**, 432–440. (doi:10.1016/j.tics.2008.07.010)

137 Hartmann, K., Goldenberg, G., Daumuller, M. & Hermsdorfer, J. 2005 It takes the whole brain to make a cup of coffee: the neuropsychology of naturalistic actions involving technical devices. *Neuropsychologia* **43**, 625–637. (doi:10.1016/j.neuropsychologia.2004.07.015)

138 Rapcsak, S. Z., Gonzalez Rothi, L. J. & Heilman, K. M. 1987 Apraxia in a patient with atypical cerebral dominance. *Brain Cogn.* **6**, 450–463. (doi:10.1016/0278-2626(87)90139-4)

139 Coppens, P., Hungerford, S., Yamaguchi, S. & Yamadori, A. 2002 Crossed aphasia: an analysis of the symptoms, their frequency, and a comparison with left-hemisphere aphasia symptomatology. *Brain Lang.* **83**, 425–463. (doi:10.1016/S0093-934X(02)00510-2)

140 Meador, K. J., Loring, D. W., Lee, K., Hughes, M., Lee, G., Nichols, M. & Heilman, K. M. 1999 Cerebral lateralization: relationship of language and ideomotor praxis. *Neurology* **53**, 2028–2031.

141 Frey, S. H., Funnell, M. G., Gerry, V. E. & Gazzaniga, M. S. 2005 A dissociation between the representation of tool-use skills and hand dominance: insights from left- and right-handed callosotomy patients. *J. Cogn. Neurosci.* **17**, 262–272. (doi:10.1162/0898929053124974)

142 Goldenberg, G. & Hagmann, S. 1998 Tool use and mechanical problem solving in apraxia. *Neuropsychologia* **36**, 581–589. (doi:10.1016/S0028-3932(97)00165-6)

143 Osiurak, F., Jarry, C., Allain, P., Aubin, G., Etcharry-Bouyx, F., Richard, I., Bernard, I. & Le Gall, D. 2009 Unusual use of objects after unilateral brain damage: the technical reasoning model. *Cortex* **45**, 769–783. (doi:10.1016/j.cortex.2008.06.013)

144 Kertesz, A. & Hooper, P. 1982 Praxis and language: the extent and variety of apraxia in aphasia. *Neuropsychologia* **20**, 275–286. (doi:10.1016/0028-3932(82)90102-6)

145 Caramazza, A. & Badecker, W. 1989 Patient classification in neuropsychological research. *Brain Cogn.* **10**, 256–295. (doi:10.1016/0278-2626(89)90056-0)

146 Iacoboni, M. & Wilson, S. M. 2006 Beyond a single area: motor control and language within a neural architecture encompassing Broca's area. *Cortex* **42**, 503–506. (doi:10.1016/S0010-9452(08)70387-3)

147 Kimura, D. 1977 Acquisition of a motor skill after left-hemisphere damage. *Brain* **100**, 527–542. (doi:10.1093/brain/100.3.527)

148 Goldenberg, G. 2001 Imitation and matching of hand and finger postures. *Neuroimage* **14**, S132–S136. (doi:10.1006/nimg.2001.0820)

149 Warren, J. E., Wise, R. J. S. & Warren, J. D. 2005 Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci.* **28**, 636–643. (doi:10.1016/j.tins.2005.09.010)

150 Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. 2009 Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19**, 2767–2796. (doi:10.1093/cercor/bhp055)

151 Damasio, A. R. & Tranel, D. 1993 Nouns and verbs are retrieved with differently distributed neural systems. *Proc. Natl Acad. Sci. USA* **90**, 4957–4960. (doi:10.1073/pnas.90.11.4957)

152 Matzig, S., Druks, J., Masterson, J. & Vigliocco, G. 2009 Noun and verb differences in picture naming: past studies and new evidence. *Cortex* **45**, 738–758. (doi:10.1016/j.cortex.2008.10.003)

153 Hillis, A. E., Oh, S. & Ken, L. 2004 Deterioration of naming nouns versus verbs in primary progressive aphasia. *Ann. Neurol.* **55**, 268–275. (doi:10.1002/ana.10812)

154 Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F. & Braun, A. R. 2009 Symbolic gestures and spoken language are processed by a common neural system. *Proc. Natl Acad. Sci. USA* **106**, 20 664–20 669. (doi:10.1073/pnas.0909197106)

155 Roy, A. C. & Arbib, M. A. 2005 The syntactic motor system. *Gesture* **1**, 7–37. (doi:10.1075/gest.5.1.03roy)

156 Allen, K., Ibara, S., Seymour, A., Cordova, N. & Botvinick, M. 2010 Abstract structural representations of goal-directed behavior. *Psychol. Sci.* **21**, 1518–1524. (doi:10.1177/0956797610383434)

157 Grodzinsky, Y. & Santi, A. 2008 The battle for Broca's region. *Trends Cogn. Sci.* **12**, 474–480. (doi:10.1016/j.tics.2008.09.001)

158 Hagoort, P. 2005 On Broca, brain, and binding: a new framework. *Trends Cogn. Sci.* **9**, 416–423. (doi:10.1016/j.tics.2005.07.004)

159 Grafton, S. T. & Hamilton, A. F. 2007 Evidence for a distributed hierarchy of action representation in the brain. *Hum. Mov. Sci.* **26**, 590–616. (doi:10.1016/j.humov.2007.05.009)

160 Koechlin, E. & Jubault, T. 2006 Broca's area and the hierarchical organization of human behavior. *Neuron* **50**, 963–974. (doi:10.1016/j.neuron.2006.05.017)

161 Goldenberg, G., Hartmann-Schmid, K., Surer, F., Daumuller, M. & Hermsdorfer, J. 2007 The impact of dysexecutive syndrome on use of tools and technical devices. *Cortex* **43**, 424–435. (doi:10.1016/S0010-9452(08)70467-2)

162 Sitnikova, T., Holcomb, P. J., Kiyonaga, K. A. & Kuperberg, G. R. 2008 Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *J. Cogn. Neurosci.* **20**, 2037–2057. (doi:10.1162/jocn.2008.20143)

163 Ahlsen, E. 2008 Embodiment in communication—aphasia, apraxia and the possible role of mirroring and imitation. *Clin. Linguist. Phon.* **22**, 311–315. (doi:10.1080/02699200801918879)

164 Goldenberg, G., Hartmann, K. & Schlott, I. 2003 Defective pantomime of object use in left brain damage: apraxia or asymbolia? *Neuropsychologia* **41**, 1565–1573. (doi:10.1016/S0028-3932(03)00120-9)

165 Hutsler, J. & Galuske, R. A. 2003 Hemispheric asymmetries in cerebral cortical networks. *Trends Neurosci.* **26**, 429–435. (doi:10.1016/S0166-2236(03)00198-X)

166 Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. & Laufs, H. 2007 Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* **56**, 1127–1134. (doi:10.1016/j.neuron.2007.09.038)

167 Arbib, M. A. 2008 From grasp to language: embodied concepts and the challenge of abstraction. *J. Physiol. Paris* **102**, 4–20. (doi:10.1016/j.jphysparis.2008.03.001)

Selection criteria

The criteria for selection are scientific excellence, originality and interest across disciplines within biology. The Editor is responsible for all editorial decisions and she makes these decisions based on the reports received from the referees and/or Editorial Board members. Many more good proposals and articles are submitted to us than we have space to print, we give preference to those that are of broad interest and of high scientific quality.

Publishing format

Phil. Trans. R. Soc. B articles are published regularly online and in print issues twice a month. Along with all Royal Society journals, we are committed to archiving and providing perpetual access. The journal also offers the facility for including Electronic Supplementary Material (ESM) to papers. Contents of the ESM might include details of methods, derivations of equations, large tables of data, DNA sequences and computer programs. However, the printed version must include enough detail to satisfy most non-specialist readers. Supplementary data up to 10Mb is placed on the Society's website free of charge. Larger datasets must be deposited in recognised public domain databases by the author.

Conditions of publication

Articles must not have been published previously, nor be under consideration for publication elsewhere. The main findings of the article should not have been reported in the mass media. Like many journals, *Phil. Trans. R. Soc. B* employs a strict embargo policy where the reporting of a scientific article by the media is embargoed until a specific time. The Editor has final authority in all matters relating to publication.

Electronic Submission details

For full submission guidelines and access to all journal content please visit the *Phil. Trans. R. Soc. B* website at **rstb.royalsocietypublishing.org**.

The Royal Society, the national academy of science of the UK and the Commonwealth, is at the cutting edge of scientific progress. We support many top young scientists, engineers and technologists, influence science policy, debate scientific issues with the public and much more. We are an independent, charitable body and derive our authoritative status from over 1400 Fellows and Foreign Members.

Our mission: to expand knowledge and further the role of science and engineering in making the world a better place.

The Royal Society's strategic priorities are to:

- invest in future scientific leaders and in innovation,
- influence policymaking with the best scientific advice,
- invigorate science and mathematics education,
- increase access to the best science internationally, and
- inspire an interest in the joy, wonder and excitement of scientific discovery.

For further information on the Society's activities, please contact the following departments on the extensions listed by dialling +44 (0)20 7839 5561, or visit the Society's Web site (royalsociety.org).

Research Support (UK grants and fellowships)
Research Appointments (Fellowships): 2542
Research Grants: 2223
International travel Grants: 2555
Newton International Fellowships: 2559

Science Policy Centre
General enquiries: 2550

Science Communication
General enquiries: 2573

Library and Information Services
Library/archive enquiries: 2606

Cover image: Misaki, a 9-year old female chimpanzee, cracking macadamia nuts with a stone hammer and anvil while carrying her 10-day old female infant. Misaki participated in the nut-cracking experiments reported by Bril *et al.* (see *Phil. Trans. R. Soc. B* (2012), **367**, 59–74). Photo taken in the outdoor enclosure of the Hayashibara Great Ape Research Institute, Japan. © Dr. Yasuko Tashiro, reproduced with permission.

From action to language: comparative perspectives on primate tool use, gesture and the evolution of human language

Papers of a Theme Issue compiled and edited by James Steele, Pier Francesco Ferrari and Leonardo Fogassi